UNIVERSIDAD NACIONAL MAYOR DE SAN MARCOS

FACULTAD DE CIENCIAS MATEMÁTICAS

E.A.P DE INVESTIGACIÓN OPERATIVA

La Dualidad en la programación no lineal

TESIS

para optar el titulo profesional de Licenciado en Investigación Operativa

AUTOR

Luis Alberto Oré Luján

Lima-Peru

1991

INTRODUCCION

El presente trabajo de Tésis para optar la Licenciatura en la Especialidad de Investigación Operativa, trata del estudio del Problema de Dualidad en el Caso No Lineal; a diferencia de la Programación Lineal donde el Problema Dual se define de manera única para cada forma de Programa Lineal, en Programación No Lineal, el Problema Primal tiene diversas formas Duales, y para cada forma Dual se establecen sus Propiedades con respecto del Problema Primal.

En el Presente trabajo se estudia el Problema Dual Minimáx; con el objeto de posibilitar el desarrollo del tema en referencia, se ha agrupado el presente trabajo en 4 Capítulos. En el Capítulo I, se establecen conceptos y propiedades matemáticas fundamentales. En el Capítulo II, se presentan algunos teoremas básicos de la Programación Matemática con sus respectivas demostraciones. En el Capítulo III, se define el concepto de Programa Matemático No Lineal y se desarrollan las condiciones de optimización de Kuhn-Tucker, así como otros teoremas, los cuales permiten caracterizar la solución óptima del Problema de Programación No Lineal. En el Capítulo IV, se define el Dual Minimáx, y se desarrollan sus propiedades con respecto del Programa Primal; el desarrollo de los temas de interés viene acompañado de sus respectivos ejemplos, lo cual hace fácil la comprensión de los mismos.

Debo finalmente dejar constancia de gratitud a nuestro asesor, Dr. Manuel Mario Renteria Vera, bajo cuya orientación se hizo parte del desarrollo del presente trabajo; con la misma intensidad y afecto debo también mi gratitud al Profesor Mg. Tomas Nuñez Lay, por su inquebrantable voluntad, para ayudarnos a desarrollar la otra parte del presente trabajo.

CONTENIDO

C A P I T U L O I

CONCEPTOS Y PROPIEDADES MATEMATICAS FUNDAMENTALES

CONCEPTOS MATEMATICOS FUNDAMENTALES

INTRODUCCION

En esta parte veremos algunos conceptos y propiedades matemáticas fundamentales, las cuales nos permitirán desarrollar las condiciones necesarias y suficientes que establecen cuando un punto X_o es *Solución Optima de un Programa Matemático.*

Tales conceptos y propiedades matemáticas son los siguientes:

01.- Función Convexa

Definición:

Sea una función: $f : S \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$, S convexo, $S \neq \emptyset$.

Se dice que f es una función convexa sobre S , si:

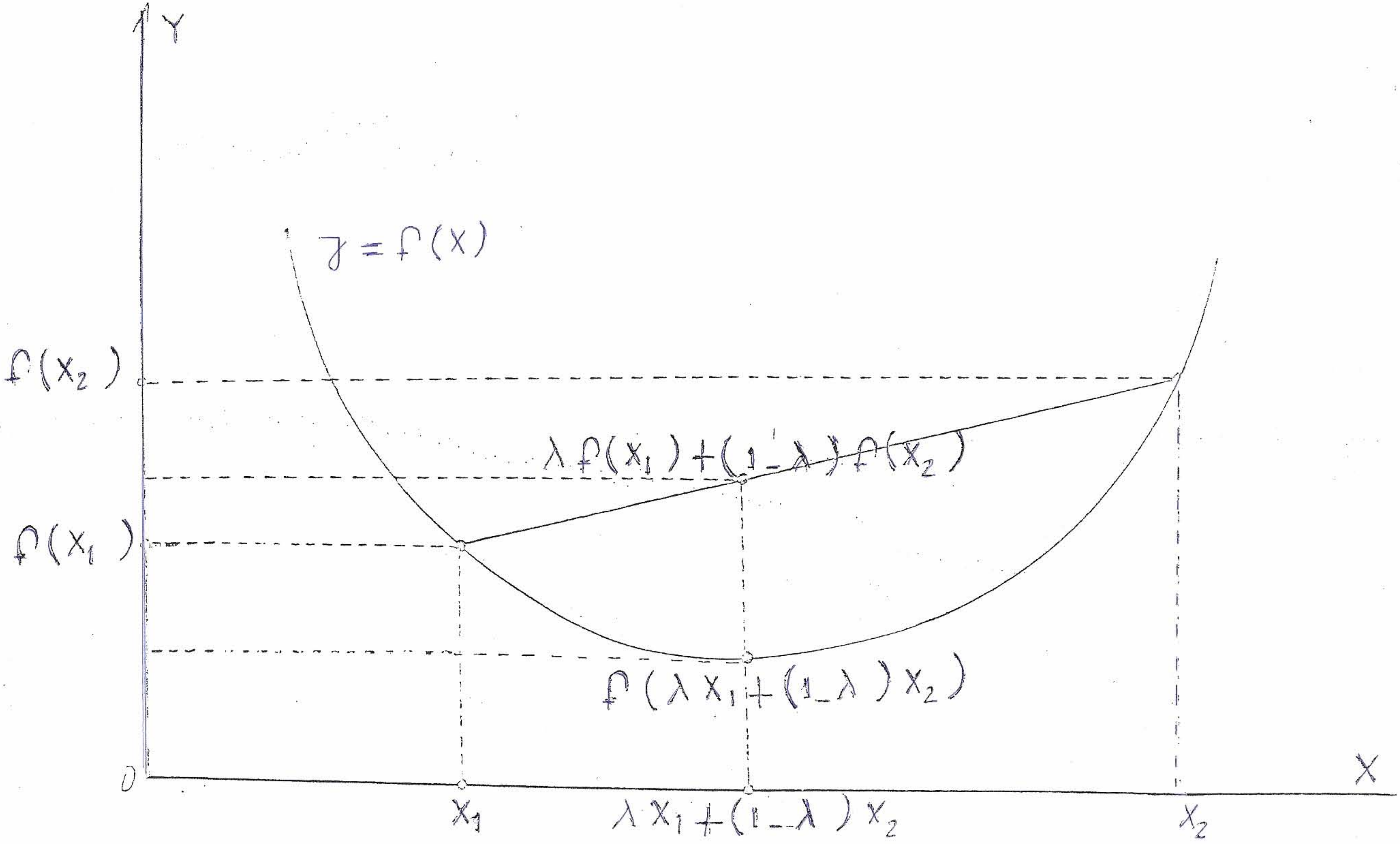
$$f(\lambda X_1 + (1-\lambda) X_2) \leq \lambda f(X_1) + (1-\lambda) f(X_2) , \forall \lambda \in [0,1] , \forall X_1,X_2 \in S .$$

Geométricamente, el concepto de *Función Convexa* se puede interpretar mediante el Gráfico No.1 ,

donde:

1) $\lambda X_1 + (1-\lambda) X_2$, $0 \leq \lambda \leq 1$, representa el conjunto de puntos del segmento de recta de extremos X_1 y X_2 .

GRAFICO No. 1


Y
$y = f(x)$
$f(x_2)$
$\lambda f(x_1) + (1 - \lambda) f(x_2)$
$f(x_1)$
$f(\lambda x_1 + (1 - \lambda) x_2)$
0
X
x_1
$\lambda x_1 + (1 - \lambda) x_2$
x_2

2) $f(\lambda X_1 + (1-\lambda) X_2)$, $0 \leq \lambda \leq 1$, representa el valor de la altura de la función en el punto $\lambda X_1 + (1-\lambda) X_2$.

3) $\lambda f(X_1) + (1-\lambda) f(X_2)$, $0 \leq \lambda \leq 1$, representa la altura de la curva que une los puntos $(X_1 , f(X_1))$ y $(X_2 , f(X_2))$ en el punto: $\lambda X_1 + (1-\lambda) X_2$, $0 \leq \lambda \leq 1$.

Y tal como puede observarse, la altura de la cuerda en el punto : $\lambda X_1 + (1-\lambda) X_2$, es al menos tan grande como la altura de la función en el punto $\lambda X_1 + (1-\lambda) X_2$, $\forall \lambda \in [0,1]$.

02.- Función Cóncava

Definición:

Sea una función: $g : S \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$, S convexo, $S \neq \emptyset$.

Se dice que g es una *Función Cóncava* sobre S , si:

$$g(\lambda X_1 + (1-\lambda) X_2) \geq \lambda g(X_1) + (1-\lambda) g(X_2) , \forall \lambda \in [0,1] , \forall X_1, X_2 \in S .$$

La interpretación geométrica es análoga al caso anterior.

03.- Matriz Semi-definida Positiva

Definición:

Sea $D \in M_{nxn} (\mathbb{R})$.

Se dice que D es una *Matriz Semi-definida Positiva*, sí y solamente sí:

$$X^t D X \geq 0 , \forall X \in M_{nx1} (\mathbb{R})$$

04.- Máximo Global

Definición:

Sea una función: $f : S \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$,

Sea $X_o \in S$.

Se dice que f tiene un *máximo global* en el punto X_o , sí y solamente sí:

$$f(X) \leq f(X_o) \quad , \quad \forall X \in S .$$

Análogamente, se define el *Mínimo Global* de una función.

05.- Mínimo Local

Definición:

Sea una función: $f : S \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$, $S \neq \emptyset$.

Sea $X_o \in S$.

Se dice que f tiene un *mínimo local* en el punto X_o , sí y solamente sí:

$$\exists \ \varepsilon > 0 \ / \ f(X_o) \leq f(X) \quad , \quad \forall \ X \in V_\varepsilon(X_o) \cap S$$

en este caso se dice que el valor $f(X_o)$ es *Mínimo Local* de f .

Análogamente, se define el *Máximo Local* de una función.

06.- Punto de Silla Local

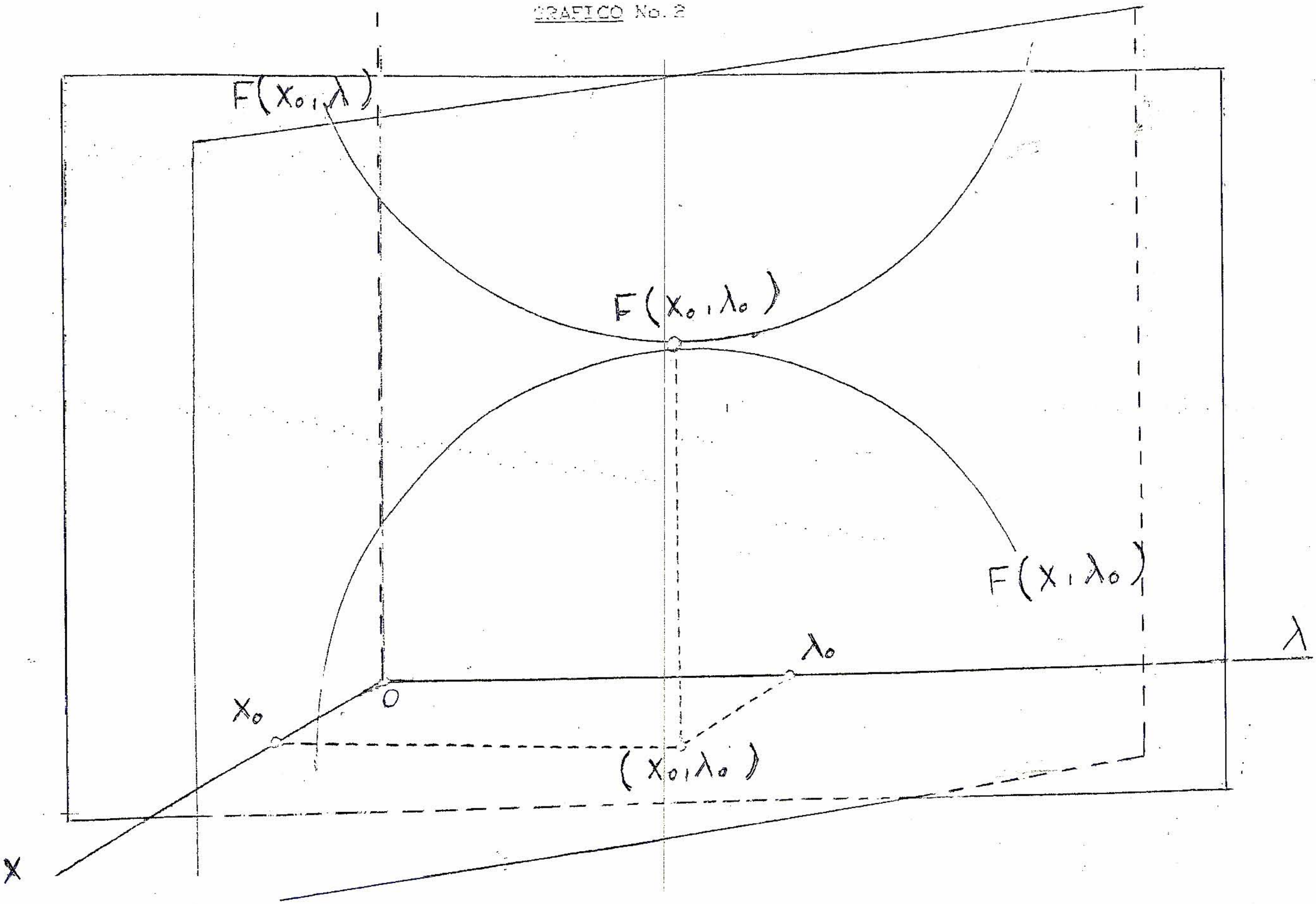
Definición:

Sea una función: $F : S_1 \times S_2 \subseteq \mathbb{R}^m \times \mathbb{R}^n \longrightarrow \mathbb{R}$, $S_1 \times S_2 \neq \emptyset$.

Sea $(X_o, \lambda_o) \in S_1 \times S_2$.

Se dice que el punto (X_o, λ_o) es punto de silla de $F(X, \lambda)$, si:

GRAFICO No. 2



$\exists\ \varepsilon > 0\ /\ \forall\ (X,\lambda) \in S_1 \times S_2$ con $\| X - X_0 \| \le \varepsilon$ y $\| \lambda - \lambda_0 \| \le \varepsilon$,

se tiene que:

1) $F(X_0,\lambda_0) \le F(X,\lambda_0)$, $\forall\ X \in S_1$. Es decir, X_0 minimiza el valor de $F(X,\lambda)$ en S_1 , para $\lambda = \lambda_0$ fijo.

2) $F(X_0,\lambda_0) \ge F(X_0,\lambda)$, $\forall\ \lambda \in S_2$. Es decir, λ_0 maximiza el valor de $F(X,\lambda)$ en S_2 , para $X = X_0$ fijo.

Conjuntamente, las 2 anteriores condiciones se escriben:

$$F(X_0,\lambda) \le F(X_0,\lambda_0) \le F(X,\lambda_0) \quad , \quad \forall\ X \in S_1 \quad , \quad \forall\ \lambda \in S_2 \quad ,$$

ó también:

$$X_0 \text{ minimiza } F(X,\lambda_0) \text{ en } S_1 \text{ ,}$$
$$\lambda_0 \text{ maximiza } F(X_0,\lambda) \text{ en } S_2 \text{ .}$$

El Gráfico No.2 ilustra geométricamente el concepto de *Punto de Silla* de la función $F(X,\lambda)$.

Notas

1) Si la definicón de *Punto de Silla* se verifica $\forall\ X \in \mathbb{R}^n$ y $\forall\ \lambda \in \mathbb{R}^m$, se dice que (X_0,λ_0) es *Punto de Silla Global* de $F(X,\lambda)$.

2) Si la definicón de *Punto de Silla* se verifica $\forall\ X \ge 0$ y $\forall\ \lambda \ge 0$, se dice que (X_0,λ_0) es un *Punto de Silla No Negativo* de $F(X,\lambda)$.

07.- Derivada Parcial

Definición:

Sea una función: $f : S \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$, S abierto, $S \neq \emptyset$.

Sean:

$$X_o = (X_1, X_2, \ldots\ldots, X_i, \ldots\ldots, X_n) \in S$$

$e_i = (0, \ldots\ldots\ldots, 1, \ldots\ldots, 0) \in S$, con todas sus componentes nulas, salvo la i-ésima componente que toma el valor 1 .

Sea $t \in \mathbb{R}$, de modo que: $X_o + t\, e_i \in S$. La *Derivada Parcial* de f en X_o , con respecto a la componente X_i , se denota con $\dfrac{\partial f(X_o)}{\partial X_i}$, y se define como:

$$\frac{\partial f(X_o)}{\partial X_i} = \lim_{t \to 0} \frac{f(X_o + t\, e_i) - f(X_o)}{t}$$

$$= \lim_{t \to 0} \frac{f(X_1, \ldots, X_{i-1}, X_i + t, X_{i+1}, \ldots, X_n) - f(X_1, \ldots, X_n)}{t}$$

si tal limite existe.

08.- Función Continua

Definición:

Sea una función: $f : S \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$, S abierto.

Sea $X_o \in S$.

Se dice que f es *Continua* en el punto X_o , si:

1) Está definida $f(X_o)$.

2) Existe $\lim_{X \to X_o} f(X)$.

3) $\lim_{X \to X_o} f(X) = f(X_o)$.

09.- El Vector Gradiente

Definición:

Sea una función: $f : S \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$ definida en el conjunto abierto S.

Sean $X_o = (X_1, X_2, \ldots\ldots, X_n) \in S$,

y supongamos que existe $\dfrac{\partial f(X_o)}{\partial X_i}$, $\forall\ i = 1,2,\ldots\ldots,n$.

Se llama *Vector Gradiente* de f en X_o , y se denota $\nabla f(X_o)$, al vector de derivadas parciales:

$$\nabla f(X_o) = \begin{bmatrix} \dfrac{\partial f(X_o)}{\partial X_1} \\ \dfrac{\partial f(X_o)}{\partial X_2} \\ \vdots \\ \dfrac{\partial f(X_o)}{\partial X_n} \end{bmatrix}_{n\times 1}$$

10.- Función Diferenciable

Definición:

Sea una función: $f : S \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$, S abierto.

Sea $X_o \in S$,

y supongamos que existe $\dfrac{\partial f(X_o)}{\partial X_i}$, $\forall\ i = 1,2,\ldots\ldots,n$.

Se dice que f es diferenciable en X_o , si existe una función:

$$g : \mathbb{R}^n \longrightarrow \mathbb{R}$$

tal que: $f(X_o + \Delta X) = f(X_o) + (\Delta X)^t\, \nabla f(X_o) + g(\Delta X)$

y $$\lim_{\|\Delta X\| \to 0} \frac{g(\Delta X)}{\|g(\Delta X)\|} = 0$$

en este caso se escribe:

$$f(X_o + \Delta X) \doteq f(X_o) + (\Delta X)^t\, \nabla f(X_o)$$

11.- Teorema

Sea una función: $f : S \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$, S abierto.

Sea $X_o \in S$.

Si $\exists\ \dfrac{\partial f}{\partial X_i}$ y es continua en $X = X_o$, $\forall\ i = 1,2,\ldots\ldots,n$.

Entonces f es diferenciable en $X = X_o$.

12.- Teorema de Taylor

Sea una función: $f : S \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$, S abierto.

Sean X_1 y $X_2 \in S$, tales que los puntos X del segmento de recta que los une: $X = \lambda X_2 + (1-\lambda) X_1$, $0 \leq \lambda \leq 1$

pertenecen a S .

Si f es diferenciable en todos los puntos de dicho segmento, entonces existe $X^* \in (X_1, X_2)$ / $f(X_2) - f(X_1) = (X_2 - X_1)^t \ \nabla f(X^*)$

13.- Función Continuamente Diferenciable

Definición:

Sea una función: $f : S \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$, S abierto.

Sea $X_0 = (X_1, X_2, \ldots\ldots, X_n) \in S$.

Se dice que f es una *Función Continuamente Diferenciable* en X_0 , si:

$$\frac{\partial f}{\partial X_i} \text{ existe y es continua en } X = X_0 \quad , \quad \forall i = 1, 2, \ldots\ldots, n .$$

14.- Teorema de Separación de Hiperplanos

Sean K_1 , $K_2 \subseteq \mathbb{R}^n$, K_1 y K_2 convexos.

Si K_1 y K_2 son disjuntos, existe un Hiperplano:

$C' X = b$, $b \neq 0$, que separa a K_1 y K_2 de la siguiente manera:

$$C' X \leq b \quad , \quad \forall X \in K_1$$
$$C' X \geq b \quad , \quad \forall X \in K_2$$

C A P I T U L O II

TEOREMAS BASICOS DE LA PROGRAMACION MATEMATICA

TEOREMAS BASICOS DE LA PROGRAMACION MATEMATICA

En esta parte veremos algunos teoremas básicos de la *Programación Matemática Convexa*, los cuales permitirán desarrollar la teoría matemática que establece las condiciones de optimización del problema de Programación Matemática.

Tales Teoremas son:

Teorema No. 1

Sea una función $f : S \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$.

S un conjunto convexo, cerrado y acotado, $S \neq \emptyset$.

Si f es una función convexa en S, entonces:

I.- f tiene a lo más un Mínimo Local.

II.- Si tal mínimo existe, este es un Mínimo Global y pertenece al conjunto convexo.

Demostración

I.- Vamos a hacer la demostración de esta parte usando el método de reducción al absurdo.

Sea X_0 un punto donde f tiene un mínimo local. Se tiene luego por definición de mínimo lócal que:

$f(X_o) \leq f(X)$, $\forall\ X \in V_\varepsilon(X_o) \cap S$, donde ε es un número real estrictamente positivo y bastante pequeño. .....(1)

Supongamos que existe otro punto $X_1 \in S$, $X_1 \neq X_o$, tal que:

$f(X_1) \leq f(X)$, $\forall\ X \in V_\delta(X_1) \cap S$, y $f(X_1) \neq f(X_o)$, es decir, un mínimo local distinto de X_o . .....(2)

Asumamos que: $f(X_1) < f(X_o)$ . .....(3)

Como X_o , $X_1 \in S$, y S es un conjunto convexo, se tiene luego que:

$$X = \lambda X_1 + (1-\lambda) X_o \in S , \quad \forall\ \lambda \in [0,1] \qquad .....(4)$$

En las relaciones (1) y (4), consideremos $\lambda_o > 0$ bastante pequeño, de modo que $X = \lambda_o X_1 + (1-\lambda_o) X_o \in V_\varepsilon(X_o)$.

Es decir:

$$\lambda_o X_1 + (1-\lambda_o) X_o \in V_\varepsilon(X_o) \cap S \qquad .....(5)$$

Se tiene luego de las relaciones (1) y (5) :

$$f(X_o) \leq f(\lambda_o X_1 + (1-\lambda_o) X_o)$$

y como por hipótesis, f es una función convexa en S :

$$\begin{aligned} f(X_o) &\leq \lambda_o f(X_1) + (1-\lambda_o) f(X_o) \\ &< \lambda_o f(X_o) + (1-\lambda_o) f(X_o) \qquad \text{.....por paso (3)} \\ &= f(X_o) \end{aligned}$$

$$\Longrightarrow \qquad f(X_o) < f(X_o) \qquad\qquad (\Longrightarrow|\Longleftarrow)$$

Como la contradicción proviene de suponer que existe más de un mínimo local, luego, debe de ocurrir lo contrario, es decir f debe de tener a lo más un mínimo local.

II.- Sea $X_o \in S$, el punto donde f tiene un mínimo local, luego:

$f(X_o) \le f(X)$, $\forall\ X \in V_\varepsilon(X_o) \cap S$, donde ε es un número real estrictamente positivo y bastante pequeño. .....(1)

Como $X_o \in S$, y S es un conjunto convexo:

$$Y = (1-\lambda)\, X_o + \lambda\, X \in S \ , \ \forall\ \lambda \in [0,1] \ , \ \forall\ X \in S \qquad .....(2)$$

de las relaciones (1) y (2), consideremos, $\lambda_1 > 0$, lo suficientemente pequeño, de modo que:

$$Y = (1-\lambda_1)\, X_o + \lambda_1\, X \in V_\varepsilon(X_o) \ , \ \forall\ X \in S.$$

Luego:

$$(1-\lambda_1)\, X_o + \lambda_1\, X \in V_\varepsilon(X_o) \cap S \ , \ \forall\ X \in S \ , \qquad .....(3)$$

y de las relaciones (1) y (3):

$$f(X_o) \le f((1-\lambda_1)\, X_o + \lambda_1\, X) \ , \ (1-\lambda_1)\, X_o + \lambda_1\, X \in V_\varepsilon(X_o) \cap S \ , \ \forall\ X \in S$$

y como por hipótesis, f es una función convexa en S:

$$\begin{aligned} f(X_o) &\le (1-\lambda_1)\, f(X_o) + \lambda_1\, f(X) \ , \quad \forall\ X \in S \\ &= f(X_o) - \lambda_1\, f(X_o) + \lambda_1\, f(X) \ , \quad \forall\ X \in S \end{aligned}$$

ó $\lambda_1 f(X_0) \le \lambda_1 f(X)$, $\lambda_1 > 0$, $\forall X \in S$

ó $f(X_0) \le f(X)$, $\forall X \in S$

luego:

X_0 es mínimo global de f . ■

Teorema No. 2

Sea una función: $f : S \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$.

S un conjunto convexo, cerrado y acotado, $S \neq \emptyset$.

Si f es una función cóncava en S , entonces:

I.- f tiene a lo más un máximo local.

II.- Si tal máximo existe, éste es un máximo global y pertenece al conjunto convexo.

Demostración

Es análoga al caso anterior. ■

Teorema No. 3

Sean las funciones: $f_i : S \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$, $i = 1,2,\ldots\ldots,m$.

S un conjunto convexo.

Si f_i es una función convexa en S , $\forall i = 1,2,\ldots\ldots,m$.

$\Longrightarrow \sum_{i=1}^{m} \lambda_i f_i$ es función convexa en S , $\lambda_i \ge 0$, $i = 1,2,\ldots\ldots,m$.

Demostración

f_i es función convexa en S , $\forall\, i = 1,2,\dots\dots,m.$

$f_i(\theta X_1 + (1-\theta) X_2) \le \theta f_i(X_1) + (1-\theta) f_i(X_2)$, $\forall\, \theta \in [0,1]$, $\forall\, X_1, X_2 \in S$,

$\forall\, i = 1,2,\dots\dots,m.$

multiplicando miembro a miembro de esta última relación por $\lambda_i \ge 0$, $i = 1,2,$ $\dots\dots, m$, se tiene:

$$\lambda_1 f_1(\theta X_1 + (1-\theta) X_2) \le \theta \lambda_1 f_1(X_1) + (1-\theta) \lambda_1 f_1(X_2)$$

$$\lambda_2 f_2(\theta X_1 + (1-\theta) X_2) \le \theta \lambda_2 f_2(X_1) + (1-\theta) \lambda_2 f_2(X_2)$$

$$\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots$$

$$\lambda_m f_m(\theta X_1 + (1-\theta) X_2) \le \theta \lambda_m f_m(X_1) + (1-\theta) \lambda_m f_m(X_2)$$

sumando miembro a miembro estas m relaciones se tiene:

$$\sum_{i=1}^{m} \lambda_i f_i(\theta X_1 + (1-\theta) X_2) \le \theta \sum_{i=1}^{m} \lambda_i f_i(X_1) + (1-\theta) \sum_{i=1}^{m} \lambda_i f_i(X_2)$$

$\forall\, \theta \in [0,1]$, $\forall\, X_1, X_2 \in S$.

$\Longrightarrow$ $\sum_{i=1}^{m} \lambda_i f_i$ es función convexa en S . ■

Teorema No. 4

Sea una función: $f : S \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$.

Sea S un conjunto convexo, $S \neq \emptyset$.
f diferenciable en S .

f es convexa en $S \Longleftrightarrow f(X_1) - f(X_2) \geq (X_1 - X_2)^t \nabla f(X_2)$, $\forall X_1, X_2 \in S$.

<u>Demostración</u>

$\Longrightarrow$) <u>Necesidad</u>

Como f es convexa en S , luego:

$$f(\lambda X_1 + (1-\lambda) X_2) \leq \lambda f(X_1) + (1-\lambda) f(X_2) \quad , \forall X_1, X_2 \in S \ , \forall \lambda \in (0,1]$$

ó

$$f(\lambda X_1 + (1-\lambda) X_2) - f(X_2) \leq \lambda \{f(X_1) - f(X_2)\} \ , \forall X_1, X_2 \in S \ , \forall \lambda \in (0,1]$$

ó

$$\frac{f(X_2 + \lambda (X_1 - X_2)) - f(X_2)}{\lambda} \leq f(X_1) - f(X_2) \ , \forall X_1, X_2 \in S \ , \forall \lambda \in (0,1]$$

luego:

$$\lim_{\lambda \to 0^+} \frac{f(X_2 + \lambda (X_1 - X_2)) - f(X_2)}{\lambda} \leq f(X_1) - f(X_2) \quad , \forall X_1, X_2 \in S \quad ..(1)$$

Por otro lado, S es convexo, entonces:

$$\lambda X_1 + (1-\lambda) X_2 \in S \ , \ \forall X_1, X_2 \in S \ , \ \forall \lambda \in (0,1]$$

además f es diferenciable en S , luego:

$$f(\lambda X_1 + (1-\lambda) X_2) \doteq f(X_2) + \lambda (X_1 - X_2)^t \nabla f(\theta)$$

para algún $\theta \in (X_2 , X_2 + \lambda (X_1 - X_2))$, $\forall X \in S$, $\forall \lambda \in (0,1]$

ó

$$\frac{f(\lambda X_1 + (1-\lambda) X_2) - f(X_2)}{\lambda} = (X_1 - X_2)^t \nabla f(\theta) .$$

para algún $\theta \in (X_2 , X_2 + \lambda (X_1 - X_2))$, $\forall X \in S$, $\forall \lambda \in (0,1]$.

Luego:

$$\lim_{\lambda \to 0^+} \frac{f(\lambda X_1 + (1-\lambda) X_2) - f(X_2)}{\lambda} = \lim_{\lambda \to 0^+} (X_1 - X_2)^t \nabla f(\theta) .$$

para algún $\theta \in (X_2 , X_2 + \lambda (X_1 - X_2))$, $\forall X_1, X_2 \in S$.

Y como ∇f es continua en S :

$$\lim_{\lambda \to 0^+} \frac{f(\lambda X_1 + (1-\lambda) X_2) - f(X_2)}{\lambda} = (X_1 - X_2)^t \nabla f \left[\lim_{\lambda \to 0^+} \theta \right] .$$

para algún $\theta \in (X_2 , X_2 + \lambda (X_1 - X_2))$, $\forall X_1, X_2 \in S$.

Luego:

$$\lim_{\lambda \to 0^+} \frac{f(\lambda X_1 + (1-\lambda) X_2) - f(X_2)}{\lambda} = (X_1 - X_2)^t \nabla f(X_2) \quad .....(2)$$

Y de las relaciones (1) y (2):

$$(X_1 - X_2)^t \nabla f(X_2) \leq f(X_1) - f(X_2)$$

(<===) Suficiencia

Por hipótesis se tiene que:

$$f(X_1) - f(X_2) \geq (X_1 - X_2)^t \nabla f(X_2) \quad , \quad \forall X_1, X_2 \in S .$$

Y vamos a demostrar que:

f es convexa en S .

Sean $X_1, X_2 \in S$, como S es un conjunto convexo:

$$X_3 = \lambda X_1 + (1-\lambda) X_2 \in S , \quad \forall \lambda \in [0,1] \qquad .....(1)$$

como X_1 y $X_3 \in S$, entonces:

$$f(X_1) - f(X_3) \geq (X_1 - X_3)^t \nabla f(X_3) , \quad \forall \lambda \in [0,1] \qquad .....(2)$$

como X_2 y $X_3 \in S$, luego:

$$f(X_2) - f(X_3) \geq (X_2 - X_3)^t \nabla f(X_3) , \quad \forall \lambda \in [0,1] \qquad .....(3)$$

Multiplicando la relación (2) por λ , la relación (3) por $(1-\lambda)$, y sumando miembro a miembro se tiene:

$$\lambda f(X_1) - \lambda f(X_3) + (1-\lambda) f(X_2) - (1-\lambda) f(X_3) \geq \lambda (X_1 - X_3)^t \nabla f(X_3) + (1-\lambda) (X_2 - X_3)^t \nabla f(X_3) \quad , \quad \forall \lambda \in [0,1]$$

ó

$$\lambda f(X_1) + (1-\lambda) f(X_2) \geq f(X_3) + \left\{ \lambda (X_1 - X_3)^t + (1-\lambda) (X_2 - X_3)^t \right\} \nabla f(X_3) ,$$

$\forall \lambda \in [0,1]$

ó

$$\lambda\, f(X_1) + (1-\lambda)\, f(X_2) \geq f(X_3) + \left\{ \lambda\, X_1^t + (1-\lambda)\, X_2^t - X_3^t \right\} \nabla f(X_3) \,.$$

$\forall\, \lambda \in [0,1]$

ó

$$\lambda\, f(X_1) + (1-\lambda)\, f(X_2) \geq f(X_3) + \left\{ \lambda\, X_1 + (1-\lambda)\, X_2 - X_3 \right\}^t \nabla f(X_3) \,.$$

$\forall\, \lambda \in [0,1]$

Y como por la relación (1): $X_3 = \lambda\, X_1 + (1-\lambda)\, X_2$.

Luego:

$$\lambda\, f(X_1) + (1-\lambda)\, f(X_2) \geq f(\lambda\, X_1 + (1-\lambda)\, X_2) \,, \quad \forall\, \lambda \in [0,1] \quad , \quad \forall\, X_1, X_2 \in S$$

es decir:

f es convexa en S. ■

<u>Teorema</u> No. 5

Sea una función: $f : S \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$.

Sea S un conjunto convexo, $S \neq \emptyset$.

f convexa y continuamente diferenciable sobre el conjunto S . $X_0 \in S$.

Se tiene entonces que:

$$f(X_0) \leq f(X) \iff (X - X_0)^t\, \nabla f(X_0) \geq 0 \quad , \forall\, X \in S \,.$$

Demostración

===>) Necesidad

$X_o \in S$, y S es un conjunto convexo, luego:

$$\lambda X + (1-\lambda) X_o \in S , \quad \forall X \in S , \forall \lambda \in (0,1] \qquad .....(1)$$

Por hipótesis:

$$f(X_o) \leq f(X) \quad , \forall X \in S . \qquad .....(2)$$

de las relaciones (1) y (2):

$$f(X_o) \leq f(\lambda X + (1-\lambda) X_o) \quad , \forall X \in S , \quad \forall \lambda \in (0,1]$$

ó

$$f(X_o + \lambda (X - X_o)) - f(X_o) \geq 0 \quad , \quad \forall X \in S , \quad \forall \lambda \in (0,1]$$

y

$$\frac{f(X_o + \lambda (X - X_o)) - f(X_o)}{\lambda} \geq 0 \quad , \quad \forall X \in S , \quad \forall \lambda \in (0,1] \qquad ...(3)$$

Por otro lado, S es convexo, entonces:

$$\lambda X + (1-\lambda) X_o \in S , \quad \forall X \in S , \forall \lambda \in (0,1]$$

y como f es diferenciable en S , luego:

$$f(X_o + \lambda (X - X_o)) \doteq f(X_o) + \lambda (X - X_o)^t \nabla f(\theta)$$

para algún $\theta \in (X_o , X_o + \lambda (X - X_o))$, $\forall X \in S , \forall \lambda \in (0,1]$

ó

$$\frac{f(X_o + \lambda (X - X_o)) - f(X_o)}{\lambda} \doteq (X - X_o)^t \nabla f(\theta) \ ,$$

para algún $\theta \in (X_o \ , X_o + \lambda (X - X_o))$, $\forall X \in S$, $\forall \lambda \in (0,1]$ ...(4)

De las relaciones (3) y (4):

$0 \le (X - X_o)^t \nabla f(\theta)$, para algún $\theta \in (X_o \ , X_o + \lambda (X - X_o))$, $\forall X \in S$,
$\forall \lambda \in (0,1]$.

luego:

$\lim_{\lambda \to 0^+} (X - X_o)^t \nabla f(\theta) \ge 0$, para algún $\theta \in (X_o \ , X_o + \lambda (X - X_o))$, $\forall X \in S$.

Y como ∇f es continua en S :

$(X - X_o)^t \nabla f\left[\lim_{\lambda \to 0^+} \theta\right] \ge 0$, para algún $\theta \in (X_o \ , X_o + \lambda (X - X_o))$,
$\forall X \in S$,

luego:

$$(X - X_o)^t \nabla f(X_o) \ge 0 \ , \quad \forall X \in S \ .$$

(<===) <u>Suficiencia</u>

Como f es continuamente diferenciable en $S \subseteq \mathbb{R}^n$, luego f es diferen-ciable en $S \subseteq \mathbb{R}^n$; además S es un conjunto convexo y f es una función convexa en S ; por Teorema anterior se tiene luego que:

$$f(X) - f(X_o) = (X - X_o)^t \nabla f(X_o) \quad , \quad \forall X \in S . \qquad .....(1)$$

Pero por hipótesis:

$$(X - X_o)^t \nabla f(X_o) \geq 0 \quad , \quad \forall X \in S \qquad .....(2)$$

luego de las relaciones (1) y (2):

$$f(X) - f(X_o) \geq 0 \quad , \quad \forall X \in S$$

ó :

$$f(X_o) \leq f(X) \quad , \quad \forall X \in S \qquad \blacksquare$$

C A P I T U L O III

EL PROGRAMA MATEMATICO NO LINEAL - CONDICIONES DE OPTIMIZACION

01.- Definición de Programa Matemático No Lineal.

02.- Función Lagrangeano.

03.- Condición de Existencia de Punto de Silla.

04.- Condición Necesaria y Suficiente para que un Punto (X_o, λ_o) sea Punto de Silla de un Programa Matemático (*Condiciones de Kuhn y Tucker*).

05.- Condición Suficiente para que un Punto (X_o, λ_o) sea Solución Optima de un Programa Matemático Convexo.

06.- Corolario.

01.- El Programa Matemático No Lineal

Definición:

Un Programa Matemático No Lineal consiste en encontrar un vector:

$$X = (X_1, X_2, \ldots\ldots\ldots\ldots, X_n)$$

el cual:

$$\min Z = f(X)$$

s. a.

$$g_i(X) \geq 0 \quad , \quad i = 1,2,\ldots\ldots,m$$

$$X \geq 0$$

donde:

1) f , $g_i : \mathbb{R}^n \longrightarrow \mathbb{R}$, $i = 1,2,\ldots\ldots,m$, son funciones continuas en $\mathbb{R}^n$

2) Al menos una de las funciones f , g_i , $i = 1,2,\ldots\ldots,m$, es No Lineal.

NOTAS:

1) Si la función objetivo del Programa Matemático No Lineal es de maximización, y contiene restricciones de igualdad, o restricciones de desigualdad en sentido contrario, se puede expresar en la forma anterior; no hay por lo tanto Pérdida de Generalidad al suponer que el problema no lineal sea de la forma dada.

2) Si las funciones f , g_i , $i = 1,2,\ldots\ldots,m$; son lineales, el programa matemático recibe el nombre de *Programa Lineal*.

3) Si las funciones f , g_i , $i = 1,2,\ldots\ldots,m$, son convexas en $\mathbb{R}^n$, el programa matemático no lineal recibe el nombre de *Programa Matemático No Lineal Convexo*.

02.- <u>Función Lagrangeano</u>

<u>Definición</u>:

Sea el programa matemático:

$$\begin{aligned} P \ : \ \ & \min Z = f(X) \\ & \text{s. a.} \\ & \quad g_i(X) \leq 0 \ , \quad i = 1,2,\ldots\ldots,m \\ & \quad X \geq 0 \end{aligned}$$

su función lagrangeano se denota con $L(X,\lambda)$, donde:

$$L(X,\lambda) = f(X) + \sum_{i=1}^{m} \lambda_i \, g_i(X)$$

y los coeficientes $\lambda_i \in \mathbb{R}^+$, $i = 1,2,\ldots\ldots,m$, reciben el nombre de *Multiplicadores de Lagrange*.

Empleando la notación Matricial:

$$\lambda^t = [\lambda_1 \ \lambda_2 \ \ldots\ldots \ \lambda_m]_{1\times m}$$

$$g(X) = \begin{bmatrix} g_1(X) \\ g_2(X) \\ \vdots \\ g_m(X) \end{bmatrix}_{m\times 1}$$

el problema P se puede expresar matricialmente de la siguiente manera:

$$P : \quad \min Z = f(X)$$

s. a.

$$g(X) \leq 0$$

$$X \geq 0$$

y su función lagrangeano:

$$L(X,\lambda) = f(X) + \lambda^t g(X) \quad , \quad \lambda \geq 0$$

NOTA:

La Función Lagrangeano: $L(X,\lambda) = f(X) + \sum_{i=1}^{m} \lambda_i \, g_i(X)$

contiene $n + m$ variables:

1) Las n componentes del vector X.
2) Las m variables λ_i asociadas a las m restricciones.

Definición:

Sea el Programa Matemático:

$$Q \ : \ \min Z = f(X)$$

$$\text{s. a.}$$

$$g(X) = 0$$

$$X \in \mathbb{R}^n$$

su función Lagrangeano se define como:

$$L(X,\lambda) = f(X) + \sum_{i=1}^{m} \lambda_i \, g_i(X) \quad , \quad \lambda_i \in \mathbb{R} \ , \ i = 1,2,\dots\dots,m$$

$$= f(X) + \lambda^t \, g(X) \quad , \quad \lambda \in \mathbb{R}^m \ .$$

Definición:

Sea el Programa Matemático:

$$R \ : \ \max Z = f(X)$$

$$\text{s. a.}$$

$$g(X) = 0$$

$$X \in \mathbb{R}^n$$

se puede expresar como:

$$\min -Z = -f(X)$$

$$\text{s. a.}$$

$$g(X) = 0$$

$$X \in \mathbb{R}^n$$

su función lagrangeano se escribe:

$$L(X,\lambda) = -f(X) + \lambda^t g(X) \quad , \quad \lambda \in \mathbb{R}^m .$$

03.- <u>Teorema</u> No.1 (Condición de Existencia del Punto de Silla)

Sea el programa matemático:

$$P : \min Z = f(X)$$

$$\text{s. a.}$$

$$g_i(X) \leq 0 \quad , \quad i = 1,2,\ldots\ldots,m$$

$$X \geq 0$$

tal que:

1) $f(X)$ y $g_i(X)$, $i = 1,2,\ldots\ldots,m$, son funciones convexas sobre el conjunto convexo:

$$S = \left\{ X = (X_1,X_2,\ldots\ldots,X_n) \in \mathbb{R}^n \ / \ X_j \geq 0 \ , \ \forall\, j = 1,2,\ldots\ldots,n \right\}$$

2) $\exists\ \overline{X} \geq 0$, tal que: $g_i(\overline{X}) < 0$, $\forall\, i = 1,2,\ldots\ldots,m$.

Se tiene que:

X_o es Solución Optima del Problema P , sí y sólo sí , $\exists\ \lambda_o \geq 0$, tal que para:

$$L(X,\lambda) = f(X) + \sum_{i=1}^{m} \lambda_i\, g_i(X)$$

se cumple que:

$$L(X_o,\lambda) \leq L(X_o,\lambda_o) \leq L(X,\lambda_o) \quad , \quad \forall\, X \geq 0 \quad , \quad \forall\, \lambda \geq 0$$

<u>Demostración</u>

(<===)

Por hipótesis se tiene que (X_o,λ_o) es punto de silla de $L(X,\lambda)$.

y vamos a demostrar en esta parte que X_o es solución óptima del Problema P.

Como (X_o, λ_o) es punto de silla de $L(X,\lambda)$, se tiene por definición de punto de silla que:

$$L(X_o,\lambda) \leq L(X_o,\lambda_o) \leq L(X,\lambda_o) \quad , \quad \forall X \geq 0 \quad , \quad \forall \lambda \geq 0$$

ó

$$f(X_o) + \sum_{i=1}^{m} \lambda_i \, g_i(X_o) \leq f(X_o) + \sum_{i=1}^{m} \lambda_i^o \, g_i(X_o) \leq f(X) + \sum_{i=1}^{m} \lambda_i^o \, g_i(X) \; ,$$

$$\forall X \geq 0 \quad , \quad \forall \lambda_i \geq 0 \quad , \quad \forall \lambda_i^o \geq 0 \quad , \; i = 1,2,\ldots\ldots,m \; . \qquad \ldots\ldots(1)$$

de donde:

$$f(X_o) + \sum_{i=1}^{m} \lambda_i \, g_i(X_o) \leq f(X_o) + \sum_{i=1}^{m} \lambda_i^o \, g_i(X_o) \quad , \; \lambda_i \geq 0 \; , \; \lambda_i^o \geq 0 \; , \; i =1,\ldots,m$$

ó

$$\sum_{i=1}^{m} \lambda_i \, g_i(X_o) \leq \sum_{i=1}^{m} \lambda_i^o \, g_i(X_o) \quad , \; \lambda_i \geq 0 \; , \; \lambda_i^o \geq 0, i = 1,2,\ldots,m \quad \ldots\ldots(2)$$

Procedamos ahora por el absurdo, y supongamos que para $i = k$, $1 \leq k \leq n$, ocurra que: $g_k(X_o) > 0$, osea que X_o no satisfaga las restricciones del Problema P.

De la relación (2), se tiene:

$$\sum_{\substack{i=1\\ i\neq k}}^{m} \lambda_i\ g_i(X_o)\ +\ \lambda_k\ g_k(X_o)\ \leq\ \sum_{i=1}^{m} \lambda_i^o\ g_i(X_o)\ ,\quad \lambda_i \geq 0\ ,\ i = 1,2,\ldots,m$$

ó

$$\lambda_k\ g_k(X_o)\ \leq\ \sum_{i=1}^{m} \lambda_i^o\ g_i(X_o)\ -\ \sum_{i=1}^{m} \lambda_i\ g_i(X_o)\ ,\quad \lambda_i \geq 0\ ,\ i = 1,2,\ldots,m$$

como: $g_k(X_o) > 0$, se tiene de esta última relación que:

$$\lambda_k\ \leq\ \frac{\displaystyle\sum_{i=1}^{m} \lambda_i^o\ g_i(X_o)\ -\ \sum_{\substack{i=1\\ i\neq k}}^{m} \lambda_i\ g_i(X_o)}{g_k(X_o)} \qquad \ldots\ldots(3)$$

como la relación (2) se debe de verificar $\forall\ \lambda_i \geq 0$, $i = 1,2,\ldots\ldots,m$, se puede elegir mediante la relación (3) un correspondiente $\lambda_k > 0$, $1 \leq k \leq n$, lo suficientemente grande, de modo que la relación (2) no se cumpla.

Por consiguiente: $g_i(X_o) \leq 0$, $i = 1,2,\ldots\ldots,m$. $\qquad \ldots\ldots(4)$

Como $X_o \geq 0$ satisface las restricciones del Problema P , X_o es por lo tanto *Solución Factible* del Problema P .

Resta demostrar que X_o es solución óptima del Problema P ; veamos: como la relación (2) se cumple para todo $\lambda \geq 0$, se tiene en particular para $\lambda = 0$ que:

$$\sum_{i=1}^{m} \lambda_i^o \, g_i(X_o) \geq 0 \qquad \ldots\ldots(5)$$

De la relación (4) y teniendo en cuenta que $\lambda_i^o \geq 0$, $i = 1,2,\ldots\ldots,m$, se tiene:

$$\sum_{i=1}^{m} \lambda_i^o \, g_i(X_o) \leq 0 \qquad \ldots\ldots(6)$$

de las relaciones (5) y (6):

$$\sum_{i=1}^{m} \lambda_i^o \, g_i(X_o) = 0 \qquad \ldots\ldots(7)$$

y de las relaciones (1) y (7):

$$f(X_o) \leq f(X) + \sum_{i=1}^{m} \lambda_i^o \, g_i(X) \; , \quad \forall \; X \geq 0 \qquad \ldots\ldots(8)$$

como $\lambda_i^o \geq 0$, $i = 1,2,\ldots\ldots,m$, se tiene de la relación (8), que:

$$f(X_o) \leq f(X) \; , \quad \text{si} \quad g_i(X) < 0 \; , \quad \forall \; X \geq 0 \qquad \ldots\ldots(9)$$

luego de las restricciones (4) y (9) , se tiene que:

X_o es *Solución Optima* y *Factible* del Problema P .

===>)

En esta parte tenemos como hipótesis que: X_o es solución óptima del Problema P, vamos a demostrar que existe $\lambda_0 \geq 0$, tal que:

(X_0, λ_0) es Punto de Silla de la función $L(X,\lambda)$.

Con este objeto seguiremos los siguientes pasos:

1°) Se definen 2 conjuntos de puntos K_1 y K_2, y demostraremos que son convexos.

2°) Demostraremos que K_1 y K_2 son disjuntos.

3°) Basado en los pasos 1°) y 2°), haremos uso del Teorema de Separación de Hiperplanos para afirmar que existe un Hiperplano: $A\,Y = b$ que separa los conjuntos de puntos definidos por K_1 y K_2.

4°) De la definición de K_1 y K_2, y del resultado obtenido en el paso 3°), demostraremos que la matriz A es positiva y su Primera Componente es Estrictamente Positiva.

5°) Basado en el paso 4°), justificaremos la existencia del vector $\lambda \geq 0$.

6°) Basado en el paso 5°), y haciendo uso de la hipótesis, demostraremos que (X_0, λ_0) es Punto de Silla de $L(X,\lambda)$.

El desarrollo de estos pasos es como sigue:

1°) Consideremos en el espacio $\mathbb{R}^{m+1}$, dos conjuntos de puntos K_1 y K_2, definidos de la siguiente manera:

$$K_1 = \left\{ Y = (y_0, y_1, \ldots, y_m) \;/\; f(X) \leq y_0 \,,\; g_i(X) \leq y_i \,,\; i = 1,2,\ldots,m \,, \right.$$

para al menos un $X \geq 0 \Big\}$

$$K_2 = \left\{ Y = (y_0, y_1, \ldots, y_m) \ / \ f(X_0) \leq y_0 \ , \quad y_i \ \text{[illegible]} \ , \ i = 1,2,\ldots,m \right\}$$

donde f y g_i , $i = 1,2,\ldots,m$, son funciones convexas en la regió S de $\mathbb{R}^n$.

Vamos primeramente a demostrar que K_1 es un conjunto convexo.

Sea $W = (\omega_0, \omega_1, \ldots, \omega_n) \in K_1$ . .....(1)

Por definición de K_1 : $\exists \ X_1 \geq 0 \ / \ f(X_1) \leq \omega_0$, $g_i(X_1) \leq \omega_i$,

$i = 1,2,\ldots, m$.

<u>Análogamente</u>: Sea $Z = (Z_0, Z_1, \ldots, Z_n) \in K_1$ . .....(2)

Por definición de K_1 : $\exists \ X_2 \geq 0 \ / \ f(X_2) \leq Z_0$, $g_i(X_2) \leq Z_i$,

$i = 1,2,\ldots,m$.

de las relaciones (1) y (2); y para $\lambda \in [0,1]$, se tiene que:

$$\hat{X} = \lambda X_1 + (1-\lambda) X_2 \geq 0$$

Por otro lado, como f es una función convexa:

$$f(\hat{X}) = f(\lambda X_1 + (1-\lambda) X_2) \leq \lambda f(X_1) + (1-\lambda) f(X_2)$$

$$\leq \lambda \omega_0 + (1-\lambda) Z_0 \qquad .....(3)$$

Como también: $g_i(X)$, $i = 1,2,\dots,m$, es también una función convexa; luego:

$$g_i(\hat{X}) = g_i(\lambda X_1 + (1-\lambda) X_2) \leq \lambda g_i(X_1) + (1-\lambda) g_i(X_2) \quad , \quad \forall i = 1,2,\dots,m$$

$$\leq \lambda \omega_i + (1-\lambda) Z_i \quad , \quad \forall i = 1,2,\dots,m$$

$$\dots\dots(4)$$

De (1), (2), (3) y (4) se ha demostrado que:

$$W , Z \in K_1 \implies \lambda W + (1-\lambda) Z \in K_1 \quad , \quad \forall \lambda \in [0,1]$$

$\therefore$ K_1 es un conjunto convexo.

Vamos ahora a demostrar que K_2 es también un conjunto convexo.

Sea $W = (\omega_0,\omega_1,\dots,\omega_n) \in K_2$. $\dots\dots(5)$

Por definición de K_2 : $f(X_0) > \omega_0$, $\omega_i < 0$, $i = 1,2,\dots,m$.

<u>Análogamente</u>: Sea $Z = (Z_0,Z_1,\dots,Z_n) \in K_2$. $\dots\dots(6)$

Por definición de K_2 : $f(X_0) > Z_0$, $Z_i < 0$, $i = 1,2,\dots,m$.

Consideremos: $\hat{X} = \lambda W + (1-\lambda) Z$, $\lambda \in [0,1]$

ó equivalentemente: $\hat{X}_i = \lambda \omega_i + (1-\lambda) Z_i$, $i = 1,2,\dots,m$, $\lambda \in [0,1]$.

Como: $W , Z \in K_2$, y $\lambda \in [0,1]$, se tiene de las relaciones (5) y (6) que:

$\hat{X}_i = \lambda\ \omega_i + (1-\lambda)\ Z_i < 0 \quad , \quad i = 1,2,\ldots,m \ , \quad \lambda \in [0,1] \quad \ldots\ldots(7)$

También de las relaciones (5) y (6), y para $\lambda \in [0,1]$, se tiene:

$$f(X_o) > \omega_o \qquad \Longrightarrow \qquad \lambda\ f(X_o) > \lambda\ \omega_o$$

también:

$$f(X_o) > Z_o \qquad \Longrightarrow \qquad (1-\lambda)\ f(X_o) > (1-\lambda)\ Z_o$$

sumando miembro a miembro:

$$\lambda\ f(X_o) + (1-\lambda)\ f(X_o) > \lambda\ \omega_o + (1-\lambda)\ Z_o$$

$$\Longrightarrow \qquad f(X_o) > \lambda\ \omega_o + (1-\lambda)\ Z_o \qquad \ldots\ldots(8)$$

De las relaciones (5), (6), (7) y (8), se ha demostrado que:

$W = (\omega_o, \omega_1, \ldots, \omega_n) \in K_2 \quad , \quad Z = (Z_o, Z_1, \ldots, Z_n) \in K_2 \quad \Longrightarrow$

$\lambda\ W + (1-\lambda)\ Z = (\lambda\ \omega_o + (1-\lambda)\ Z_o ,\ \lambda\ \omega_1 + (1-\lambda)\ Z_1 , \ldots, \lambda\ \omega_n + (1-\lambda)\ Z_n) \in K_2 ,$

$\forall\ \lambda \in [0,1]$.

$\Longrightarrow$ K_2 es un Conjunto Convexo.

2°) Vamos ahora a demostrar que K_1 y K_2 son conjuntos disjuntos.

Por el absurdo; supongamos que: $K_1 \cap K_2 \neq \emptyset$.

Como: $K_1 \cap K_2 \neq \emptyset \quad \Longrightarrow \quad \exists\ Y = (y_o, y_1, \ldots, y_m) \in K_1 \cap K_2$

$\Longrightarrow \quad Y \in K_1 \wedge Y \in K_2$.

Como: $Y = (y_o, y_1, \ldots, y_m) \in K_1$, luego:

a) $f(X) \leq y_o$, para algún $X \geq 0$.

b) $g_i(X) \leq y_i$, $i = 1, 2, \ldots, m$.

Análogamente: $Y \in K_2$, luego:

c) $f(X_o) > y_o$

d) $y_i < 0$, $i = 1, 2, \ldots, m$.

De las relaciones a) y c):

$f(X) \leq y_o < f(X_o) \implies f(X) < f(X_o)$, para algún $X \geq 0$

.....(1)

Como X_o es solución óptima del Problema P :

$$f(X_o) \leq f(X) \quad , \quad \forall X \geq 0 \qquad \text{.....(2)}$$

De las relaciones (1) y (2):

$$f(X_o) < f(X_o) \quad ,$$

lo cual es una contradicción; como la contradicción proviene de suponer que: $K_1 \cap K_2 \neq \emptyset$, debe de ocurrir lo contrario; es decir: $K_1 \cap K_2 = \emptyset$.

3°) Vamos a demostrar en esta parte que $\exists \lambda_o \geq 0$.
De las partes 1°) y 2°). Se tiene que: K_1 y K_2 son conjuntos convexos y disjuntos; por Teorema de Separación de Hiperplanos, se tiene

que:

$\exists$ un Hiperplano : $A\,Y = b$, $A \in M_{1xm+1}(\mathbb{R}) \neq O_{1xm+1}$, $b \in \mathbb{R}-\{0\}$,

que separa a K_1 y K_2 de la siguiente manera:

$A\,Y_1 \geq b$, $\forall\, Y_1 = (y_{01}, y_{11}, y_{21}, \ldots\ldots, y_{m1}) \in K_1$

$A\,Y_2 \geq b$, $\forall\, Y_2 = (y_{02}, y_{12}, y_{22}, \ldots\ldots, y_{m2}) \in K_2$

luego:

$$A\,Y_1 \geq A\,Y_2 \ , \ \forall\, Y_1 \in K_1 \ , \ \forall\, Y_2 \in K_2$$

ó

$$\sum_{i=0}^{m} a_i\, y_{i1} \geq \sum_{i=0}^{m} a_i\, y_{i2} \qquad \ldots\ldots(1)$$

Como: $Y_2 = (y_{02}, y_{12}, y_{22}, \ldots\ldots, y_{m2}) \in K_2$

luego: $y_{i2} > 0$, $i = 1,2,\ldots\ldots,m$,

y la relación (1) se verifica $\forall\, y_{i1} \in \mathbb{R}$, $i = 0,1,2,\ldots\ldots,m$, y $\forall\, y_{i2} \in \mathbb{R}^-$, $i = 1,2,\ldots\ldots,m$.

4°) Como: $A = (a_0, a_1, \ldots\ldots, a_{i_o}, \ldots\ldots, a_m)$, supongamos que existe i_o fijo,

$0 \leq i_o \leq m \ / \ a_{i_o} < 0 \qquad \ldots\ldots(1)$

De la relación (1) del paso anterior:

$$\sum_{i=0}^{m} a_i\, y_{i1} \geq \sum_{\substack{i=0 \\ i\neq i_o}}^{m} a_i\, y_{i2} + a_{i_o}\, y_{i_o 2}$$

ó

$$a_{i_o}\, y_{i_o 2} \leq \sum_{i=0}^{m} a_i\, y_{i1} - \sum_{\substack{i=0 \\ i\neq i_o}}^{m} a_i\, y_{i2}$$

y teniendo acá en cuenta la relación (1) anterior:

$$y_{i_o 2} \geq \frac{\displaystyle\sum_{i=0}^{m} a_i\, y_{i1} - \sum_{\substack{i=0 \\ i\neq i_o}}^{m} a_i\, y_{i2}}{a_{i_o}} \qquad \ldots\ldots(2)$$

luego, la relación (1) de paso 3°) se verifica para vectores Y_2, tales que su componente $y_{i_o 2}$ satisfaga la relación (2).

Por otro lado, como $Y_2 \in K_2$, se puede elegir un correspondiente vector Y_2, cuya i_o-ésima componente, $0 \leq i_o \leq m$, sea tan pequeña como sea posible, de tal manera que la relación (2) no se cumpla.

Concluimos entonces que debe de ser: $A = (a_0, a_1, \ldots, a_{i_o}, \ldots, a_m) \geq 0$.

Consideremos 2 vectores:

$$Y_1 = \left[f(X) \,,\, g_1(X) \,,\, \ldots\ldots \,,\, g_m(X) \right] \in K_1 \qquad \ldots\ldots(3)$$

$$Y_2 = \left[f(X_o) \,,\, 0 \,,\, \ldots\ldots \,,\, 0 \right]$$

Como la relación (1) del paso 3°) se verifica también para los puntos que se encuentran en la Frontera del conjunto K_2, se tiene de la relación (1) del paso 3°) y de la relación (3):

$$a_o\, f(X) + a_1\, g_1(X) + \ldots\ldots + a_m\, g_m(X) \geq a_o\, f(X_o) \quad , \quad \forall\, X \geq 0 \;,$$

$$\forall\, a_i \geq 0 \quad , \; i = 1,2,\ldots\ldots,m \;. \qquad \ldots\ldots(4)$$

Vamos ahora a demostrar que a_o es estrictamente positiva.

Por el absurdo: Supongamos $a_o = 0$, en la relación (4) se tiene entonces:

$$a_1\, g_1(X) + a_2\, g_2(X) + \ldots\ldots + a_m\, g_m(X) \geq 0 \;, \; \forall\, X \geq 0 \;, \; \forall\, a_i \geq 0 \;,$$

$$i = 1,2,\ldots\ldots,m \;. \qquad \ldots\ldots(5)$$

Como por hipótesis existe algún $X \geq 0$, digamos $\overline{X}$, para el cual:

$$g_i(\overline{X}) < 0 \;, \; i = 1,2,\ldots\ldots,m \;; \quad y \;:\; a_i \geq 0 \;, \; i = 1,2,\ldots\ldots,m \;,$$

luego:

$$a_1\, g_1(\overline{X}) + a_2\, g_2(\overline{X}) + \ldots\ldots + a_m\, g_m(\overline{X}) < 0 \qquad \ldots\ldots(6)$$

De las relaciones (5) y (6) se sigue que:

$$a_1\, g_1(\bar{X}) + \ldots\ldots + a_m\, g_m(\bar{X}) \;<\; a_1\, g_1(\bar{X}) + \ldots\ldots + a_m\, g_m(\bar{X}) \;,$$

lo cual es una contradicción.

Como la contradicción proviene de suponer que $a_o = 0$, y es $a_o \geq 0$, debe por lo tanto de ser: $a_o > 0$.

5°) Veamos ahora la existencia del vector: $\lambda_o = (\lambda_1^o, \lambda_2^o, \ldots\ldots, \lambda_m^o) \geq 0.$

Como $a_o > 0$, dividiendo miembro a miembro la relación (4) por a_o ,se tiene:

$$f(X) + \frac{a_1}{a_o}\, g_1(X) + \ldots\ldots + \frac{a_m}{a_o}\, g_m(X) \;\geq\; f(X_o) \;,\; \forall\, X \geq 0 \;,\; \forall\, a_i \geq 0 \;,$$

$i = 1,2,\ldots\ldots,m$ . .....(1)

Haciendo $\lambda_i^o = \dfrac{a_i}{a_o} \geq 0$, $i = 1,2,\ldots\ldots,m$, queda justificada la existencia del vector λ .

6°) En la relación (1) se tiene:

$$f(X) + \lambda_1^o\, g_1(X) + \ldots\ldots + \lambda_m^o\, g_m(X) \;\geq\; f(X_o) \;,\; \forall\, X \geq 0 \;,\; \forall\, \lambda_i^o \geq 0 \;,$$

$i = 1,2,\ldots\ldots,m$.

En particular, para: $X = X_o \geq 0$:

$$f(X_o) + \lambda_1^o\, g_1(X_o) + \ldots\ldots + \lambda_m^o\, g_m(X_o) \;\geq\; f(X_o) \;.$$

Como: $L(X_o,\lambda_o) = f(X_o) + \lambda_1^o g_1(X_o) + \ldots\ldots + \lambda_m^o g_m(X_o)$

.....(1)

se tiene de estas 2 últimas relaciones que:

$$L(X_o,\lambda_o) \geq f(X_o) \qquad \ldots\ldots(2)$$

Por otro lado, de la relación (6) se tiene:

$$\lambda_1^o g_1(X_o) + \lambda_2^o g_2(X\ ) + \ldots\ldots + \lambda_m^o g_m(X_o) \geq 0 \qquad \ldots\ldots(3)$$

y como por hipótesis es:

$$g_i(X_o) \leq 0 \quad \text{y} \quad \lambda_i^o \geq 0 \quad , \quad \forall\ i = 1,2,\ldots\ldots,m .$$

Luego:

$$\lambda_1^o g_1(X_o) + \lambda_2^o g_2(X_o) + \ldots\ldots + \lambda_m^o g_m(X_o) \leq 0 \qquad \ldots\ldots(4)$$

de las relaciones (3) y (4):

$$\lambda_1^o g_1(X_o) + \lambda_2^o g_2(X\ ) + \ldots\ldots + \lambda_m^o g_m(X_o) = 0 \qquad \ldots\ldots(5)$$

y de las relaciones (1) y (5):

$$L(X_o,\lambda_o) = f(X_o) \qquad \ldots\ldots(6)$$

Por otro lado, como:

$$g_i(X_o) \leq 0 \quad , \quad \text{para} \quad \lambda_i \geq 0 \quad , \quad i = 1,2,\ldots\ldots,m .$$

se tiene:

$$\sum_{i=1}^{m} \lambda_i \; g_i(X_o) \leq 0$$

luego:

$$f(X_o) \geq f(X_o) + \sum_{i=1}^{m} \lambda_i \; g_i(X_o)$$

ó

$$f(X_o) \geq L(X_o, \lambda) \quad , \quad \forall \lambda \geq 0 \qquad \ldots\ldots(7)$$

Por otra parte se tiene de la relación (1) del paso 5°), y de la relación (5):

$$L(X, \lambda_o) = f(X) + \lambda_1^o \; g_1(X) + \lambda_2^o \; g_2(X) + \ldots\ldots + \lambda_m^o \; g_m(X) \geq f(X_o)$$

$$\geq f(X_o) + 0$$

$$\geq f(X_o) + \lambda_1^o \; g_1(X_o) + \lambda_2^o \; g_2(X_o) + \ldots\ldots + \lambda_m^o \; g_m(X_o) = L(X_o, \lambda_o)$$

$$\Longrightarrow L(X, \lambda_o) \geq L(X_o, \lambda_o) \quad , \quad \forall X \geq 0 \qquad \ldots\ldots(8)$$

y de las relaciones (6), (7) y (8):

$$L(X_o, \lambda) \leq L(X_o, \lambda_o) \leq L(X, \lambda_o) \quad , \quad \forall X \geq 0 \; , \; \forall \lambda \geq 0 \; .$$

$\Longrightarrow$ (X_o, λ_o) es Punto de Silla de $L(X, \lambda)$. ■

EJEMPLO: Se sabe que: $(X_0, \lambda_0) = (X_1^0 , X_2^0 , \lambda_1^0 , \lambda_2^0)$

$$= \left(\frac{1}{2} , \frac{3}{2} , 0 , 1 \right)$$

es Punto de Silla del Lagrangeano del Programa Matemático:

$$\min Z = (X_1 - 1)^2 + (X_2 - 2)^2$$

s. a.

$$-X_1 + X_2 \leq 1$$

$$X_1 + X_2 \leq 2$$

$$X_1 \geq 0$$

$$X_2 \geq 0$$

usando el Teorema de Existencia de Punto de Silla, ¿ Podría determinar su solución óptima ?

SOLUCION:

Se tiene que:

1) La función objetivo y las restricciones son funciones convexas en $\mathbb{R}^2$, en particular lo son en el 1er. Cuadrante.

2) $\exists$ $X = (X_1, X_2) \geq 0$, digamos por ejemplo: $(X_1, X_2) = (0, 1/2)$, tal que;

para : $g_1(X_1, X_2) = - X_1 + X_2 - 1$

y : $g_2(X_1, X_2) = X_1 + X_2 - 2$

se tiene:

$$g_1(0,1/2) = -0 + \frac{1}{2} - 1 = -\frac{1}{2} < 0$$

$$g_2(0,1/2) = 0 + \frac{1}{2} - 2 = -\frac{3}{2} < 0$$

De la hipótesis del Problema, y de las condiciones 1) y 2), se tiene por Teorema de existencia de punto de silla que:

$X_0 = (X_1^0, X_2^0) = \left(\frac{1}{2}, \frac{3}{2}\right)$ es solución óptima del Programa Matemático dado.

04.- Teorema No. 2 (Condiciones de Kuhn-Tucker)

Sea el Programa Matemático No Lineal:

$$P : \min Z = f(X)$$

s. a.

$$g_i(X) \leq 0 \quad , \quad i = 1,2,\ldots\ldots,m$$

$$X \geq 0$$

donde: $f \,,\, g_i : S \subset \mathbb{R}^n \longrightarrow \mathbb{R} \,,\; i = 1,2,\ldots\ldots,m$.

Son funciones convexas y continuamente diferenciables en el conjunto:

$$S = \left\{ X = (X_1,X_2,\ldots\ldots,X_n) \in \mathbb{R}^n \;/\; X_i \geq 0 \quad , \; i = 1,2,\ldots\ldots,n \right\}$$

el punto (X_o,λ_o) / a) $X_o \in S$

b) $\lambda_o \geq 0$

es Punto de Silla del Lagrangeano $L(X,\lambda)$:

$$L(X,\lambda) = f(X) + \sum_{i=1}^{m} \lambda_i \, g_i(X)$$

sí y sólo sí:

$$\text{I.-} \quad \frac{\partial L(X,\lambda)}{\partial X_j}\bigg|_{(X_o,\lambda_o)} = \frac{\partial f(X)}{\partial X_j}\bigg|_{X_o} + \sum_{i=1}^{m} \lambda_i \frac{\partial g_i(X)}{\partial X_j}\bigg|_{(X_o,\lambda_o)} \geq 0 \,,$$

$\forall\, j = 1,2,\ldots\ldots,n$.

$$\text{II.-}\quad (X^o)^t \frac{\partial L(X_o,\lambda_o)}{\partial X} = \sum_{j=1}^{n} X_j^o \left\{ \frac{\partial f(X_o)}{\partial X_j} + \sum_{i=1}^{m} \lambda_i^o \frac{\partial g_i(X_o)}{\partial X_j} \right\} = 0 .$$

$$\text{III.-}\quad \left.\frac{\partial L(X,\lambda)}{\partial \lambda_i}\right|_{(X_o,\lambda_o)} = g_i(X_o) \le 0 \quad , \quad \forall\ i = 1,2,\ldots\ldots,m$$

$$\text{IV.-}\quad \lambda_o^t \frac{\partial L(X_o,\lambda_o)}{\partial \lambda} = \sum_{i=1}^{m} \lambda_i^o\, g_i(X_o) = 0$$

<u>Demostración</u>

===>) <u>Condición Necesaria</u>

Vamos a demostrar en esta parte que, si (X_o,λ_o) es punto de silla de $L(X,\lambda)$, luego (X_o,λ_o) satisface las condiciones de Kuhn-Tucker.

I) Como (X_o,λ_o) es Punto de Silla de $L(X,\lambda)$, se tiene por definición de Punto de Silla que:

$$L(X_o,\lambda_o) \le L(X,\lambda_o) \quad , \quad \forall\ X \ge 0 .$$

En particular, para $X = X_o + h\, e_j \ge 0$, $\forall\ j = 1,2,\ldots,n$, y $\forall\ h > 0$ se tiene:

$$L(X_o,\lambda_o) \le L(X_o + h\, e_j \,,\, \lambda_o) \quad , \quad \forall\ h > 0 \quad , \quad \forall\ j = 1,2,\ldots\ldots,n .$$

Luego:

$$L(X_o + h\,e_j\,,\,\lambda_o) - L(X_o,\lambda_o) \;\geq\; 0 \quad , \quad \forall\, h > 0 \; ,$$

y :

$$\frac{L(X_o + h\,e_j\,,\,\lambda_o) - L(X_o,\lambda_o)}{h} \;\geq\; 0 \quad , \quad \forall\, h > 0 \; ,$$

Como $L(X,\lambda)$ es diferenciable en (X_o,λ_o):

$$\exists \lim_{h \to o^+} \frac{L(X_o + h\,e_j\,,\,\lambda_o) - L(X_o,\lambda_o)}{h} \;\geq\; 0 \quad , \quad \forall\, j = 1,2,\ldots\ldots,n$$

Luego:

$$\frac{\partial\, L(X_o,\lambda_o)}{\partial\, X_j} \;\geq\; 0 \quad , \quad \forall\, j = 1,2,\ldots\ldots,n \; ,$$

con lo cual se verifica la 1ra. condición.

II) Vamos ahora a demostrar la 2da. condición de Kuhn-Tucker:

$$(X^o)^t \; \frac{\partial\, L(X_o,\lambda_o)}{\partial\, X} \; = \; 0 \, .$$

Por hipótesis es $X_o = (\, X^o_1, X^o_2, \ldots, X^o_j, \ldots, X^o_n \,) \;\geq\; 0$.

Es decir:

$$X^o_j \;\geq\; 0 \quad , \quad \forall\, j = 1,2,\ldots\ldots,n \; ,$$

Teniendo en cuenta este resultado y la condición I) anterior, la condición II) equivale a demostrar el siguiente resultado:

$$X_j^o \frac{\partial L(X_o, \lambda_o)}{\partial X_j} = 0 \quad , \quad \forall j = 1,2,\dots\dots,n$$

Como: $X_j^o \geq 0$, $\forall j = 1,2,\dots\dots,n$, esto implica 2 posibilidades:

1) $X_j^o = 0$ ó 2) $X_j^o > 0$.

Veamos:

1) Sea: $1 \leq j \leq n$; si $X_j^o = 0$, es claro que:

$$X_j^o \frac{\partial L(X_o, \lambda_o)}{\partial X_j} = 0 \quad , \quad \text{para } 1 \leq j \leq n .$$

2) Sea: $1 \leq j \leq n$; si $X_j^o > 0$, debe de ser: $\dfrac{\partial L(X_o, \lambda_o)}{\partial X_j} = 0$.

Por el absurdo, supongamos que: $\dfrac{\partial L(X_o, \lambda_o)}{\partial X_j} \neq 0$, para algún $j = 1,2,\dots\dots,n$.

Como de la condición I) anterior es:

$$\frac{\partial L(X_o, \lambda_o)}{\partial X_j} \geq 0 \quad , \quad \forall j = 1,2,\dots\dots,n .$$

Resulta de estas 2 últimas hipótesis:

$$\frac{\partial\, L(X_o,\lambda_o)}{\partial\, X_j} > 0 \quad , \quad \text{para algún } j = 1,2,\ldots\ldots,n$$

y como por otro lado, $L(X,\lambda)$ es continuamente diferenciable en (X_o,λ_o) , se tiene considerando la anterior relación que:

$$\exists \lim_{h\to o} \frac{L(X_o + h\, e_j\, ,\, \lambda_o) - L(X_o,\lambda_o)}{h} > 0 \quad , \text{ para algún } j = 1,2,\ldots,n\ .$$

y por Teorema de existencia del Límite Lateral, $\exists\ h < 0$ tal que:

$$\frac{L(X_o + h\, e_j\, ,\, \lambda_o) - L(X_o,\lambda_o)}{h} > 0 \quad , \text{ para algún } j = 1,2,\ldots,n\ .$$

de donde:

$$L(X_o + h\, e_j\, ,\, \lambda_o) - L(X_o,\lambda_o) < 0 \quad , \text{ para algún } j = 1,2,\ldots\ldots,n\ .$$

ó

$$L(X_o + h\, e_j\, ,\, \lambda_o) < L(X_o,\lambda_o) \quad , \text{ para algún } j = 1,2,\ldots\ldots,n\ .$$

ó

$L(\overline{X}\ ,\lambda_o) \leq L(X_o,\lambda_o)$, para: $\overline{X} = X_o + h\, e_j \geq 0$, y algún $j = 1,2,\ldots\ldots,n$.

lo cual es una contradicción , pues $L(X,\lambda_o)$ se minimiza en $X = X_o$, $\forall\ X \geq 0$.

Como la contradicción proviene de suponer que:

$$\frac{\partial L(X_o,\lambda_o)}{\partial X_j} \neq 0 \quad , \quad \text{para algún} \quad j = 1,2,\ldots\ldots,n \quad , \text{ debe de ser por}$$

lo tanto:

$$\frac{\partial L(X_o,\lambda_o)}{\partial X_j} = 0 \quad , \quad \forall\ j = 1,2,\ldots\ldots,n \ .$$

de donde:

$$X_j^o \frac{\partial L(X_o,\lambda_o)}{\partial X_j} = 0 \quad , \quad \forall \quad j = 1,2,\ldots\ldots,n \quad .$$

Finalmente de las condiciones 1) y 2) :

$$(X^o)^t \frac{\partial L(X_o,\lambda_o)}{\partial X_j} = 0$$

que es lo que se quería demostrar.

III) Vamos a demostrar en esta parte la 3ra. condición de Kuhn-Tucker:

$$\left.\frac{\partial L(X,\lambda)}{\partial \lambda_i}\right|_{(X_o,\lambda_o)} = \frac{\partial L(X_o,\lambda_o)}{\partial \lambda_i} = g_i(X_o) \leq 0 \ , \ i = 1,2,\ldots\ldots,m.$$

como:

$$L(X,\lambda) = f(X) + \sum_{i=1}^{m} \lambda_i \, g_i(X)$$

y (X_o,λ_o) es punto de silla de $L(X,\lambda)$, luego:

$$L(X_o,\lambda) \leq L(X_o,\lambda_o) \quad , \quad \forall\, \lambda \geq 0$$

ó

$$f(X_o) + \sum_{i=1}^{m} \lambda_i \, g_i(X_o) \leq f(X_o) + \sum_{i=1}^{m} \lambda_i^o \, g_i(X_o) \quad , \quad \lambda_i \geq 0 \quad , \; i = 1,2,\ldots\ldots,m$$

ó

$$\sum_{i=1}^{m} \lambda_i \, g_i(X_o) \leq \sum_{i=1}^{m} \lambda_i^o \, g_i(X_o) \quad , \quad \lambda_i \geq 0 \quad , \; i = 1,2,\ldots\ldots,m \qquad \ldots\ldots(1)$$

por otro lado:

$$\frac{\partial\, L(X_o,\lambda_o)}{\partial\, \lambda_i} = g_i(X_o) \quad , \quad i = 1,2,\ldots\ldots\ldots,m. \qquad \ldots\ldots(2)$$

y debe de demostrarse que:

$$\frac{\partial\, L(X_o,\lambda_o)}{\partial\, \lambda_i} = g_i(X_o) \leq 0 \quad , \quad \forall\, i = 1,2,\ldots\ldots\ldots,m. \qquad \ldots\ldots(3)$$

Procedamos por el absurdo, es decir, supongamos que:

$$\frac{\partial L(X_o,\lambda_o)}{\partial \lambda_i} = g_i(X_o) \le 0 \ , \quad \text{para algún } i = 1,2,\ldots\ldots,m \ ,$$

digamos i_o , $1 \le i_o \le m$, tal que:

$$g_{i_o}(X_o) > 0 \qquad \ldots\ldots(4)$$

de la relación (1) se tiene:

$$\sum_{\substack{i=1\\ i\ne i_o}}^{m} \lambda_i\, g_i(X_o) + \lambda_{i_o} g_{i_o}(X_o) \le \sum_{\substack{i=1\\ i\ne i_o}}^{m} \lambda_i\, g_i(X_o) \ , \quad \lambda_i \ge 0 \ , \ i = 1,2,\ldots\ldots,m$$

ó

$$\lambda_{i_o} g_{i_o}(X_o) \le \sum_{i=1}^{m} \lambda_i^o\, g_i(X_o) - \sum_{\substack{i=1\\ i\ne i_o}}^{m} \lambda_i\, g_i(X_o) \ , \quad \lambda_i \ge 0 \ , \ i = 1,2,\ldots\ldots,m$$

teniendo en cuenta la relación (4) en esta última relación:

$$\lambda_{i_o} \ge \frac{\displaystyle\sum_{i=1}^{m} \lambda_i^o\, g_i(X_o) - \sum_{\substack{i=o\\ i\ne i_o}}^{m} \lambda_i\, g_i(X_o)}{g_{i_o}(X_o)} \ , \quad \text{para algún } i_o \ ,$$

$$1 \le i_o \le m \qquad \ldots\ldots(5)$$

lo cual es una contradicción; ya que de la relación (1) : $\lambda_i \ge 0$,

$\forall\, i = 1,2,\ldots\ldots,m$.

Y podríamos seleccionar acá λ_{i_o} tan grande como se quiera, de modo que la relación (5) no se verifique.

Como la contradicción proviene de suponer que: $\dfrac{\partial L(X_o,\lambda_o)}{\partial \lambda_i} \leq 0$, para algún $i = 1,2,\ldots\ldots,m$, debe por lo tanto de ser:

$$\frac{\partial L(X_o,\lambda_o)}{\partial \lambda_i} = g_i(X_o) \leq 0 \quad , \quad \forall\, i = 1,2,\ldots\ldots,m \ .$$

IV) En esta parte demostraremos la 4ta. condición de Kuhn-Tucker:

$$\lambda_o^t \frac{\partial L(X_o,\lambda_o)}{\partial \lambda} = \sum_{i=1}^{m} \lambda_i^o\, g_i(X_o) = 0$$

como (X_o,λ_o) es punto de silla de $L(X,\lambda)$, luego por Teorema anterior:

$$\lambda_o^t \frac{\partial L(X_o,\lambda_o)}{\partial \lambda_i} = \sum_{i=1}^{m} \lambda_i^o\, g_i(X_o) = 0 \ .$$

(<===) Condición Suficiente

En esta parte demostraremos que si (X_o,λ_o) satiaface las condiciones de Kuhn-Tucker, entonces (X_o,λ_o) es punto de silla de $L(X,\lambda)$.

Tenemos por hipótesis que : f, $g_i : S \subset \mathbb{R}^n \longrightarrow \mathbb{R}$, $i = 1,2,\ldots,m$, son funciones convexas en S ; como $L(X,\lambda)$ es combinación lineal positiva de funciones convexas en S , luego:

$$L(X,\lambda) = f(X) + \sum_{i=1}^{m} \lambda_i \, g_i(X) \quad \text{es función convexa en } S , \ \forall \ X \in S ,$$

$\forall \ \lambda_i \geq 0$, $i = 1,2,\ldots,m$ .....(1)

En particular:

$L(X,\lambda_o)$ es una función convexa en S , para $\lambda = \lambda_o \geq 0$ fijo.

Además:

$L(X,\lambda_o)$ es continuamente diferenciable en S.

De estos dos últimos resultados, y del Teorema No. 4 del Capítulo II:

$$L(X,\lambda_o) - L(X_o,\lambda_o) \geq (X - X_o)^t \frac{\partial L(X_o,\lambda_o)}{\partial X} , \quad \forall \ X \in S$$

ó

$$L(X,\lambda_o) \geq L(X_o,\lambda_o) + X^t \frac{\partial L(X_o,\lambda_o)}{\partial X} - X_o^t \frac{\partial L(X_o,\lambda_o)}{\partial X} , \quad \forall \ X \in S$$

.....(2)

como (X_o,λ_o) satisface las condiciones de Kuhn-Tucker, se tiene por hipótesis I :

$$\frac{\partial L(X_o,\lambda_o)}{\partial X_j} \geq 0 \quad , \quad \forall \ j = 1,2,\ldots,m$$

ó

$$\frac{\partial\, L(X_o, \lambda_o)}{\partial\, X} \geq 0$$

y como $X \in S$, se tiene de esta última relación:

$$X^t \frac{\partial\, L(X_o, \lambda_o)}{\partial\, X} \geq 0 \qquad \ldots\ldots(3)$$

reemplazando los resultados de la hipótesis II , y la relación (3) en la relación (2);
se tiene:

$$L(X, \lambda_o) \geq L(X_o, \lambda_o) \quad , \quad \forall\ X \in S \qquad \ldots\ldots(4)$$

Por otro lado de la relación (1) . Se tiene en particular para $X = X_o$, que:

$L(X_o, \lambda)$ es una función convexa $\forall\ \lambda \geq 0$.

Análogamente:

$L(X_o, \lambda)$ es continuamente diferenciable $\forall\ \lambda \geq 0$ y lineal en λ .

de estas 2 últimas relaciones, y del Teorema No. 4 del Capítulo II :

$$L(X_o, \lambda) - L(X_o, \lambda_o) = (\lambda - \lambda_o)^t \frac{\partial\, L(X_o, \lambda_o)}{\partial\, X}$$

$$L(X_o,\lambda) = L(X_o,\lambda_o) + (\lambda - \lambda_o)^t \frac{\partial L(X_o,\lambda_o)}{\partial X}$$

$$= L(X_o,\lambda_o) + \lambda^t \frac{\partial L(X_o,\lambda_o)}{\partial X} - \lambda_o^t \frac{\partial L(X_o,\lambda_o)}{\partial X} \qquad \ldots\ldots(5)$$

Por hipótesis III, como (X_o,λ_o) satisface las condiciones de Kuhn-Tucker:

$$\frac{\partial L(X_o,\lambda_o)}{\partial \lambda_i} \leq 0 \quad , \quad \forall\ i = 1,2,\ldots\ldots,m\,.$$

ó

$$\frac{\partial L(X_o,\lambda_o)}{\partial \lambda} \leq 0$$

y como $\lambda \geq 0$, se tiene de esta última relación:

$$\lambda^t \frac{\partial L(X_o,\lambda_o)}{\partial \lambda} \leq 0 \qquad \ldots\ldots(6)$$

reemplazando los resultados de la hipótesis IV y de la relacióm (6) en la relación (5); se tiene:

$$L(X_o,\lambda) \leq L(X_o,\lambda_o) \quad , \quad \forall\ \lambda \geq 0 \qquad \ldots\ldots(7)$$

de las relaciones (4) y (7):

$$L(X_o,\lambda) \leq L(X_o,\lambda_o) \leq L(X,\lambda_o) \quad , \quad \forall\ X \geq 0 \quad , \quad \forall\ \lambda \geq 0$$

luego, (X_o,λ_o) es punto de silla de $L(X,\lambda)$. ∎

NOTAS:

1) De la condición IV de Kuhn-Tucker:

$$\lambda_o^t \frac{\partial L(X_o,\lambda_o)}{\partial \lambda} = \sum_{i=1}^{m} \lambda_i^o g_i(X_o) = 0 \iff \lambda_i^o g_i(X_o) = 0 \ , \ \forall i = 1,2,\dots,m$$

ya que todos los sumandos: $\lambda_i^o g_i(X_o)$, $\forall \ i = 1,2,\dots\dots,m$.

son del mismo signo: $\lambda_i^o \geq 0$ y $g_i(X_o) \leq 0$, $\forall \ i = 1,2,\dots\dots,m$.

2) Del resultado: $\lambda_i^o g_i(X_o) = 0$, $\forall \ i = 1,2,\dots\dots,m$,

se establece que:

a) Si la i-ésima restricción no esta saturada; es decir, si: $g_i(X_o) < 0$, para algún $i = 1,2,\dots\dots,m$; su multiplicador asociado λ_i es nulo:

$$g_i(X_o) < 0 \implies \lambda_i = 0 \ , \ \text{para algún } i = 1,2,\dots\dots,m .$$

b) Si la i-ésima restricción esta saturada; es decir, si: $g_i(X_o) = 0$, para algún $i = 1,2,\dots\dots,m$; su multiplicador λ_i asociado es no negativo:

$$g_i(X_o) = 0 \implies \lambda_i \geq 0 \ , \ \text{para algún } i = 1,2,\dots\dots,m .$$

EJEMPLO No. 1:

Verifique si el Punto: $(X_o,\lambda_o) = (X_1^o , X_2^o , \lambda_1^o , \lambda_2^o)$

$$= \left(\frac{1}{2} , \frac{3}{2} , 0 , 1 \right)$$

es Punto de Silla del Lagrangeano del Programa :

$$\min Z = (X_1 - 1)^2 + (X_2 - 2)^2$$

s. a.

$$-X_1 + X_2 \leq 1$$
$$X_1 + X_2 \leq 2$$
$$X_1 \geq 0$$
$$X_2 \geq 0$$

SOLUCION:

Vamos a hacer la verificación respectiva usando las condiciones de Kuhn-Tucker Tenemos primeramente un Programa Matemático Convexo, donde la función objetivo y las restricciones son funciones continuamente diferenciables en $\mathbb{R}^2$, en particular lo son en el 1er. Cuadrante.

Tenemos ahora la función Lagrangeano:

$$L(X,\lambda) = (X_1 - 1)^2 + (X_2 - 2)^2 + \lambda_1 (-X_1 + X_2 - 1) + \lambda_2 (X_1 + X_2 - 2)$$

donde:

1) $$\frac{\partial L(X,\lambda)}{\partial X_1} = 2(X_1 - 1) - \lambda_1 + \lambda_2$$

$$\frac{\partial L(X,\lambda)}{\partial X_2} = 2(X_1 - 1) + \lambda_1 + \lambda_2$$

y se observa que :

$$\frac{\partial L(X_o,\lambda_o)}{\partial X_1} = \frac{\partial L(1/2\ ,\ 3/2\ ,\ 0\ ,\ 1)}{\partial X_1} = 2\ (1/2 - 1) - 0 + 1 = 0 \geq 0$$

$$\frac{\partial L(X_o,\lambda_o)}{\partial X_2} = \frac{\partial L(1/2\ ,\ 3/2\ ,\ 0\ ,\ 1)}{\partial X_2} = 2\ (3/2 - 2) + 0 + 1 = 0 \geq 0$$

es decir, se cumple la 1ra. condición de Kuhn-Tucker.

2) $$X_o^t \frac{\partial L(X_o,\lambda_o)}{\partial X} = \sum_{j=1}^{2} X_j^o \frac{\partial L(X_o,\lambda_o)}{\partial X_j}$$

$$= X_1^o \frac{\partial L(X_o,\lambda_o)}{\partial X_1} + X_2^o \frac{\partial L(X_o,\lambda_o)}{\partial X_2}$$

$$= (1/2)\ (0) + (3/2)\ (0)$$

$$= 0$$

y se cumple la 2da. condición de Kuhn-Tucker.

3) $$\frac{\partial L(X,\lambda)}{\partial \lambda_1} = - X_1 + X_2 - 1$$

$$\frac{\partial L(X,\lambda)}{\partial \lambda_2} = X_1 + X_2 - 2$$

y se tiene que:

$$\frac{\partial L(X_o,\lambda_o)}{\partial \lambda_1} = \frac{\partial L(1/2\ ,\ 3/2\ ,\ 0\ ,\ 1)}{\partial \lambda_1} = -1/2 + 3/2 - 1 = 0 \leq 0$$

$$\frac{\partial L(X_o,\lambda_o)}{\partial \lambda_2} = \frac{\partial L(1/2\ ,\ 3/2\ ,\ 0\ ,\ 1)}{\partial \lambda_2} = 1/2 + 3/2 - 2 = 0 \leq 0$$

y se cumple la 3ra. condición de Kuhn-Tuker.

4) $$\lambda_o \frac{\partial L(X_o,\lambda_o)}{\partial \lambda} = \sum_{i=1}^{2} \lambda_i^o \frac{\partial L(X_o,\lambda_o)}{\partial \lambda_i}$$

$$= \lambda_1^o \frac{\partial L(X_o,\lambda_o)}{\partial \lambda_1} + \lambda_2^o \frac{\partial L(X_o,\lambda_o)}{\partial \lambda_2}$$

$$= (0)\quad(0) \quad + (1)\quad(0)$$

$$= 0$$

y se cumple la 4ta. condición de Kuhn-Tucker.

De las condiciones 1) , 2) , 3) y 4) :

$(X_o,\lambda_o) = \left(\frac{1}{2}\ ,\ \frac{3}{2}\ ,\ 0\ ,\ 1\right)$ es Punto de Silla del Lagrangeano del Programa Matemático dado.

EJEMPLO No. 2:

Dado el Programa Matemático:

$$\min Z = -(X-1)^3$$

s. a.

$$X \leq 2$$

$$X \geq 0$$

1) Hallar su solución óptima.

2) Usando las condiciones de Kuhn-Tucker ¿ Es posible asegurar la Optimalidad de la solución hallada en el paso anterior ?

SOLUCION

1) Hagamos $f(X) = -(X-1)^3 \quad , \quad 0 \leq X \leq 2$

Sean $X_1 \, , \, X_2 \in [0,2] \quad / \quad X_1 < X_2 \quad \Longrightarrow \quad X_1 - 1 < X_2 - 1$

$$\Longrightarrow \quad (X_1 - 1)^3 < (X_2 - 1)^3$$

$$\Longrightarrow \quad f(X_2) = -(X_2 - 1)^3 < -(X_1 - 1)^3 = f(X_1)$$

$\Longrightarrow$ f es decreciente en $[0,2]$ $\Longrightarrow$ f toma su *Valor Mínimo* en $X_o = 2$;

y

$$\min Z = \min \; -(X-1)^3 = -(2-1)^3 = -1 \, .$$

Otra manera de hallar el Mínimo:

$$f(X) = -(X-1)^3$$

$$f'(X) = -3(X-1)^2 \leq 0 \quad , \quad \forall \; X \in [0,2]$$

$\Longrightarrow$ $f'(X)$ es decreciente en $[0,2]$ $\Longrightarrow$ f toma su *Valor Mínimo* en $X_o = 2$

$\Longrightarrow$ *min* $-(X-1)^3 = -(2-1)^3 = -1$

2) Tenemos que el Lagrangeano del Programa Matemático es:

$$L(X,\lambda) = -(X-1)^3 + \lambda\,(X-2)$$

y las condiciones de Kuhn-Tucker son:

a) $$\frac{\partial L(X_o,\lambda_o)}{\partial X} = -3\,(X_o - 1)^2 + \lambda_o \geq 0$$

b) $$X_o\,\frac{\partial L(X_o,\lambda_o)}{\partial X} = -3\,(X_o - 1)^2\,X_o + \lambda_o\,X_o \geq 0$$

c) $$\frac{\partial L(X_o,\lambda_o)}{\partial \lambda} = X_o - 2 \leq 0$$

d) $$\lambda_o\,\frac{\partial L(X_o,\lambda_o)}{\partial \lambda} = \lambda_o\,(X_o - 2) \leq 0$$

Si consideramos el Vector $(X_o,\lambda_o) = (1,0)$, vemos claramente que se satisfacen las condiciones de Kuhn-Tucker , pero no podemos asegurar que $X_o = 1$ sea solución óptima del Programa dado, debido a que la función objetivo no es convexa.

05- <u>Teorema</u> No.3 (Condición Suficiente para que un Punto (X_o, λ_o) sea Solución Optima de un Programa Matemático Convexo)

Sea el Programa Matemático :

$$P : \min Z = f(X)$$

$$\text{s. a.}$$

$$g_i(X) \leq 0 \quad , \quad i = 1,2,\ldots\ldots,m$$

$$X \geq 0$$

donde: $f \ , \ g_i : S \subset \mathbb{R}^n \longrightarrow \mathbb{R}$

son funciones convexas y continuamente diferenciables en el conjunto:

$$S = \left\{ X = (X_1, X_2, \ldots, X_n) \in \mathbb{R}^n \ / \ X_i \geq 0 \quad , \ i = 1,2,\ldots,n \right\}$$

si el punto (X_o, λ_o) satisface las condiciones de Kuhn-Tucker, entonces X_o es solución óptima del problema P .

<u>Demostración</u>

Tenemos como hipótesis que (X_o, λ_o) satisface las condiciones de Kuhn-Tucker; vamos a demostrar que X_o es Solución Optima del Problema P .

Se sabe que: $L(X,\lambda) = f(X) + \lambda^t g(X)$ es función convexa, por ser combinación lineal no negativa de las funciones convexas $f(X)$ y $g(X)$, $\forall X \in S$, $\forall \lambda \geq 0$.

En particular, para $\lambda = \lambda_o$ fijo:

$$L(X,\lambda_o) = f(X) + \lambda_o^t g(X) \text{ es función convexa } \forall X \in S \quad \ldots\ldots(1)$$

como $\lambda = \lambda_o$ fijo, hagamos:

$$\tilde{f}(X) = L(X, \lambda_o) \qquad \ldots\ldots(2)$$

De las relaciones (1) y (2) :

$$\tilde{f}(X) = f(X) + \lambda_o^t\, g(X) \quad , \quad \forall\, X \in S$$

$$= f(X) + \sum_{i=1}^{m} \lambda_i^o\, g_i(X) \quad , \quad \forall\, X \in S \qquad \ldots\ldots(3)$$

Por hipótesis:

$$g_i(X) \leq 0 \quad y \quad \lambda_i^o \geq 0 \quad , \quad \forall\, i = 1,2,\ldots\ldots,m \ .$$

Luego:

$$\sum_{i=1}^{m} \lambda_i^o\, g_i(X) \leq 0 \quad , \quad \forall\, X \in S$$

y

$$f(X) + \sum_{i=1}^{m} \lambda_i^o\, g_i(X) \leq f(X) \quad , \quad \forall\, X \in S \qquad \ldots\ldots(4)$$

de las relaciones (3) y (4) :

$$\tilde{f}(X) \leq f(X) \quad , \quad \forall\, X \in S \qquad \ldots\ldots(5)$$

por otro lado:

$$\tilde{f}(X_o) = L(X_o, \lambda_o) = f(X_o) + \sum_{i=1}^{m} \lambda_i^o\, g_i(X_o)$$

$$= f(X_o)$$ ..... Por hipótesis y Condición IV del Teorema anterior.

Luego:

$$\tilde{f}(X_o) = f(X_o) \qquad \text{.....(8)}$$

De las relaciones (1) y (2); se tiene que, como $\tilde{f}(X)$ es función convexa y continuamente diferenciable, $\forall X \in S$, $\forall \lambda \geq 0$, luego por Teorema No. 4 del Capitulo II:

$$\tilde{f}(X) - \tilde{f}(X_o) \geq \nabla \tilde{f}(X_o)(X - X_o)^t \quad , \quad \forall X \in S$$

lo que por la relación (2) se puede expresar:

$$L(X,\lambda_o) - L(X_o,\lambda_o) \geq \frac{\partial L(X_o,\lambda_o)}{\partial X}(X - X_o)^t \quad , \quad \forall X \in S$$

$$= \frac{\partial L(X_o,\lambda_o)}{\partial X} X^t - \frac{\partial L(X_o,\lambda_o)}{\partial X} X_o^t \quad , \quad \forall X \in S$$

Como por hipótesis y Condición I del Teorema anterior:

$$\frac{\partial L(X_o,\lambda_o)}{\partial X} \geq 0 \qquad \text{y} \qquad X^t \geq 0$$

y por hipótesis y Condición II del Teorema anterior:

$$\frac{\partial\, L(X_o, \lambda_o)}{\partial\, X}\, X_o^t = 0$$

luego:

$$L(X, \lambda_o) - L(X_o, \lambda_o) \geq 0 \quad , \quad \forall\, X \in S$$

ó :

$$L(X_o, \lambda_o) \leq L(X, \lambda_o) \quad , \quad \forall\, X \in S$$

ó :

$$\tilde{f}(X_o) \leq \tilde{f}(X) \quad , \quad \forall\, X \in S \qquad \ldots\ldots(7)$$

De las relaciones (6) , (7) y (5):

$$f(X_o) = \tilde{f}(X_o) \leq \tilde{f}(X) \leq f(X) \quad , \quad \forall\, X \in S$$

luego :

$$f(X_o) \leq f(X) \quad , \quad \forall\, X \in S$$

es decir:

X_o es Solución Optima del Problema P . ■

NOTAS:

1) De los dos últimos Teoremas que acabamos de demostrar, se establece que las condiciones de Kuhn-Tucker son condiciones necesarias y sufi-

cientes de optimalidad global, para el caso de un Programa Matemático Convexo No Lineal.

2) En general, las condiciones de Kuhn-Tucker son condiciones necesarias y no suficientes de optimalidad global.

3) En el caso de Programas Convexos y Diferenciables, las condiciones de Kuhn-Tucker y de Punto de Silla del Lagrangeano coinciden, y no se diferencian mas que en la forma utilizada para enunciar las mismas condiciones necesarias y suficientes de optimalidad global.

Por lo tanto, en tales condiciones de diferenciabilidad y convexidad, coinciden los conjuntos de:

a) Soluciones Optimas de Programa.

b) Punto de Silla del Lagrangeano.

c) Soluciones de las Condiciones de Kuhn-Tucker.

<u>EJEMPLO</u>: Sabiendo que:

$$(X_0, \lambda_0) = (X_1^0, X_2^0, \lambda_1^0, \lambda_2^0) = \left(\frac{1}{2}, \frac{3}{2}, 0, 1 \right)$$

es Punto de Silla del Lagrangeano del Programa Matemático:

$$\text{min } Z = (X_1 - 1)^2 + (X_2 - 2)^2$$

s. a.

$$- X_1 + X_2 \leq 1$$

$$X_1 + X_2 \leq 2$$

$$X_1 \geq 0$$

$$X_2 \geq 0$$

¿ Podria determinar su Solución Optima ?

SOLUCION

Como el Programa Matemático dado es convexo, y la función objetivo y las restricciones son funciones continuamente diferenciables en $\mathbb{R}^2$, en particular lo son en el 1er. Cuadrante, se tiene luego por el Teorema que acabamos de ver que:

$X_0 = (X_1^0, X_2^0) = (1/2 , 3/2)$ es solución óptima del Programa Matemático dado.

08. - Corolario

Con las mismas condiciones del Teorema anterior, si (X_o,λ_o) es Punto de Silla de $L(X,\lambda)$ $\Longrightarrow$ $L(X_o,\lambda_o) = f(X_o)$.

Demostración

Como (X_o,λ_o) es Punto de Silla de $L(X,\lambda) = f(X) + \sum_{i=1}^{m} \lambda_i\, g_i(X)$,

se tiene por la IV condición de Kuhn-Tucker:

$$L(X_o,\lambda_o) = f(X_o) + \sum_{i=1}^{m} \lambda_i^o\, g_i(X_o) = f(X_o)$$

Luego:

$$L(X_o,\lambda_o) = f(X_o) \qquad \blacksquare$$

NOTA:

La búsqueda del óptimo de Programa Matemático Convexo:

$$P : \min Z = f(X)$$

s. a.

$$g_i(X) \leq 0 \;, \quad i = 1,2,\ldots\ldots,m$$

$$X \in S$$

quedará resuelto, si es posible encontrar un Punto de Silla (X_o,λ_o) de su

Lagrangeano; ya que si (X_0, λ_0) es Punto de Silla de su Lagrangeano, el Vector X_0 es Optimo Global del Programa.

C A P I T U L O IV

UNA FORMA DUAL NO LINEAL : EL DUAL MINIMAX

01.- Introducción.

02.- Definición de Función Dual.

03.- Teorema de Concavidad de la Función Dual.

04.- Teorema de Relación entre los Valores Optimos de la Función Dual y de la Función Objetivo del Programa Matemático Primal.

05.- Teorema de Diferenciabilidad de la Función Dual.

06.- Definición de Programa Dual Minimáx.

07.- Teorema de Relación Fundamental entre los Programas Primal y Dual.

08.- Aplicaciones.

UNA FORMA DUAL NO LINEAL: EL DUAL MINIMAX

.- Introducción

Tal como se puede establecer de los teoremas anteriores, el Problema de hallar el óptimo de un Programa Matemático queda resuelto, si es posible encontar un Punto de Silla de su Lagrangeano; ya que si (X_o,λ_o) es Punto de Silla del Lagrangeano, el vector X_o es óptimo global del Programa Matemático.

El Problema de búsqueda del Punto de Silla del Lagrangeano, consiste en minimizar el Lagrangeano con respecto del vector X, suponiendo que el vector λ se fija en su valor óptimo $\lambda = \lambda_o$.

Es decir:

$$L(X_o,\lambda_o) = \min_{X \in S} L(X,\lambda_o)$$

La dificultad asociada con esta minimización está en que no se conoce a priori el vector óptimo λ_o de los multiplicadores λ, lo cual permitiría sustituirlo en $L(X,\lambda)$, y de esta manera se obtendría una función dependiente solamente del vector X; para resolver esta dificultad, se estudiará el comportamiento del mínimo de $L(X,\lambda)$ con respecto a X, en función de los valores de los multiplicadores λ; es decir, en vez de considerar un único valor λ_o de los multiplicadores λ, que generalmente no es conocido a Priori, estudiaremos el valor del mínimo del Lagrangeano con respecto a X, para todos los $\lambda \geq 0$, osea, como función de los valores de los multiplicadores λ.

Con este objeto se define previamente la función h, denominada *Función Dual* del Programa considerado, como aquella función que hace corresponder a cada vector de multiplicadores $\lambda \geq 0$, el mínimo del Lagrangeano $L(X,\lambda)$ con respecto a $X \in S$; esto es:

$$h(\lambda) = \min_{X \in S} L(X,\lambda)$$

a partir de este concepto se define la *Forma Dual Minimáx* del Programa Matemático No Lineal:

$$\max_{\lambda \in D \subset \mathbb{R}^m} h(\lambda)$$

y luego se establecen sus propiedades con respecto al Programa Matemático considerado, al cual se le denomina *Problema Primal*.

02.- La Función Dual

Definición

Sea el Programa Matemático: Q : $\min f(X)$

s. a.

$$g_i(X) \leq 0 \quad , \; i = 1,2,\ldots\ldots,m$$

$$X \in S \subset \mathbb{R}^n_+$$

se denomina Función Dual del Programa Q , a la función:

$$h : \mathbb{R}^m \longrightarrow \mathbb{R}$$
$$\lambda \longrightarrow h(\lambda) = \min_{X \in S} L(X,\lambda)$$

y su dominio admisible D , es el conjunto de valores del vector $\lambda \geq 0$, para los cuales existe dicho mínimo:

$$D = \left\{\lambda \in \mathbb{R}^m \ / \ \lambda \geq 0 \ \text{ y existe } \min_{X \in S} L(X,\lambda)\right\}$$

<u>NOTAS</u>:

1) El conjunto D puede ser vacío; en este caso no existe ningún vector $\lambda \geq 0$, para el cual el Lagrangeano tiene un mínimo finito con respecto a $X \in S$.

2) El vector X_o que minimiza el Lagrangeano en S para un λ dado $(\lambda \neq \lambda_o)$ no necesariamente verifica las condiciones:

$$g_i(X_o) \leq 0 \quad , \quad i = 1,2,\ldots\ldots,m .$$

3) El mínimo del Lagrangeano con respecto a $X \in S$, no necesariamente existe para todos los vectores $\lambda \geq 0$; obsérvese que D es el conjunto de vectores $\lambda \geq 0$, para los cuales dicho mínimo existe; su naturaleza depende de las propiedades de la función $L(X,\lambda)$, y del conjunto S .

Ejemplos:

1) Determine la función dual, y el dominio de definición de la función dual, para el siguiente Programa Matemático:

$$Q: \quad \min Z = (X_1 - 1)^2 + (X_2 - 2)^2$$

s. a.

$$X_1 + X_2 \leq 2$$
$$X_1 \geq 0$$
$$X_2 \geq 0$$

Solución:

El Lagrangeano del Problema Q es:

$$L(X_1, X_2, \lambda) = (X_1 - 1)^2 + (X_2 - 2)^2 + \lambda (X_1 + X_2 - 2) \quad , \lambda \geq 0 .$$

su función dual:

$$h(\lambda) = \min_{\substack{X_1 \geq 0 \\ X_2 \geq 0}} L(X_1, X_2, \lambda)$$

$$= \min_{\substack{X_1 \geq 0 \\ X_2 \geq 0}} \left[(X_1 - 1)^2 + (X_2 - 2)^2 + \lambda (X_1 + X_2 - 2)\right]$$

$$= \min_{X_1 \geq 0} \left[(X_1 - 1)^2 + \lambda X_1\right] + \min_{X_2 \geq 0} \left[(X_2 - 2)^2 + \lambda X_2\right] - 2\lambda$$

hallemos ahora:

1) $\min\limits_{X_1 \geq 0} \left[(X_1 - 1)^2 + \lambda X_1 \right]$

como:

$$\frac{\partial \left[(X_1 - 1)^2 + \lambda X_1 \right]}{\partial X_1} = 2 (X_1 - 1) + \lambda = 0 \quad ===> \quad X_1 = 1 - \frac{\lambda}{2}$$

$$\frac{\partial^2 \left[(X_1 - 1)^2 + \lambda X_1 \right]}{\partial X_1^2} = 2 > 0$$

luego: $X_1 = 1 - \frac{\lambda}{2}$ es mínimo de $Y = (X_1 - 1)^2 + \lambda X_1$.

y como: $X_1 \geq 0$, $X_1 = 1 - \frac{\lambda}{2}$, entonces: $0 \leq \lambda \leq 2$.

por lo tanto:

$$X_1^0 = X_1^0(\lambda) = \begin{cases} 1 - \frac{\lambda}{2} & , \text{ si } \quad 0 \leq \lambda \leq 2 \\ 0 & \text{ si } \quad \lambda > 2 \end{cases}$$

2) $\min\limits_{X_2 \geq 0} \left[(X_2 - 2)^2 + \lambda X_2 \right]$

como:

$$\frac{\partial \left[(X_2 - 2)^2 + \lambda X_2\right]}{\partial X_2} = 2 (X_2 - 2) + \lambda = 0 \quad ===> \quad X_2 = 2 - \frac{\lambda}{2}$$

$$\frac{\partial^2 \left[(X_2 - 2)^2 + \lambda X_2\right]}{\partial X_2^2} = 2 > 0$$

luego: $X_2 = 2 - \frac{\lambda}{2}$ es mínimo de $Y = (X_2 - 2)^2 + \lambda X_2$.

y como: $X_2 \geq 0$, $X_2 = 2 - \frac{\lambda}{2}$, entonces: $0 \leq \lambda \leq 4$.

por lo tanto:

$$X_2^o = X_2^o(\lambda) = \begin{cases} 2 - \frac{\lambda}{2} & , \text{ si } 0 \leq \lambda \leq 4 \\ 0 & \text{ si } \lambda > 4 \end{cases}$$

3) Hallando ahora la Función Dual $h(\lambda)$:

a) $0 \leq \lambda \leq 2$: $X_1^o = X_1^o(\lambda) = 1 - \frac{\lambda}{2}$

$$X_2^o = X_2^o(\lambda) = 2 - \frac{\lambda}{2}$$

se tiene:

$$h(\lambda) = (X_1^o - 1)^2 + (X_2^o - 2)^2 + \lambda\,(X_1^o + X_2^o - 2)$$

$$= (1 - \frac{\lambda}{2} - 1)^2 + (2 - \frac{\lambda}{2} - 2)^2 + \lambda\,(1 - \frac{\lambda}{2} + 2 - \frac{\lambda}{2} - 2)$$

$$= \frac{\lambda^2}{4} + \frac{\lambda^2}{4} + \lambda\,(1 - \lambda)$$

$$= \frac{\lambda^2}{2} + \lambda - \lambda^2$$

$$= - \frac{\lambda^2}{2} + \lambda$$

b) $2 \leq \lambda \leq 4$: $X_1^o = X_1^o(\lambda) = 0$

$X_2^o = X_2^o(\lambda) = 2 - \frac{\lambda}{2}$

$$h(\lambda) = (X_1^o - 1)^2 + (X_2^o - 2)^2 + \lambda\,(X_1^o + X_2^o - 2)$$

$$= (0 - 1)^2 + (2 - \frac{\lambda}{2} - 2)^2 + \lambda\,(0 + 2 - \frac{\lambda}{2} - 2)$$

$$= 1 + \frac{\lambda^2}{4} - \frac{\lambda^2}{2}$$

$$= - \frac{\lambda^2}{4} + 1$$

c) $\lambda > 4$: $X_1^o = X_1^o(\lambda) = 0$

$$X_2^0 = X_2^0(\lambda) = 0$$

$$h(\lambda) = (X_1^0 - 1)^2 + (X_2^0 - 2)^2 + \lambda\,(X_1^0 + X_2^0 - 2)$$

$$= 1 + 4 - 2\lambda$$

$$= 5 - 2\lambda$$

y de las relaciones a) , b) y c) , se tiene la Función Dual:

$$h(\lambda) = \begin{cases} -\dfrac{\lambda^2}{2} + \lambda & 0 \leq \lambda \leq 2 \\ -\dfrac{\lambda^2}{4} + 1 & 2 \leq \lambda \leq 4 \\ 5 - 2\lambda & \lambda > 4 \end{cases}$$

y su dominio admisible es la semirecta real positiva:

$$D = \left\{ \lambda \in \mathbb{R} \;/\; \lambda \geq 0 \right\}$$

03.- Teorema No.1 (Concavidad de la Función Dual).

La Función Dual h del Programa Matemático Q es cóncava en cualquier subconjunto convexo $\hat{D}$ de su dominio de definición D ; independientemente de que si el Programa Q sea o no convexo.

Demostración:

Sea $\hat{D}$ es un subconjunto convexo de D , $\lambda_1, \lambda_2 \in \hat{D}$, y sea $\alpha \in [0,1]$ cualquiera.

Se tiene luego por definición de $h(\lambda)$:

$$h\left[\alpha \lambda_1 + (1 - \alpha) \lambda_2\right] = \min_{X \in S} L\left[X , \alpha \lambda_1 + (1 - \alpha) \lambda_2\right]$$

$$= \min_{X \in S} L\left[\alpha L(X,\lambda_1) + (1 - \alpha) L(X,\lambda_2)\right]$$

..... (Por ser $L(X,\lambda)$ Función Lineal en λ)

$$\geq \alpha \min_{X \in S} L(X,\lambda_1) + (1 - \alpha) \min_{X \in S} L(X,\lambda_2)$$

$$= \alpha h(\lambda_1) + (1 - \alpha) h(\lambda_2)$$

$$\therefore \quad h\left[\alpha \lambda_1 + (1 - \alpha) \lambda_2\right] \geq \alpha h(\lambda_1) + (1 - \alpha) h(\lambda_2) , \quad \forall \lambda_1, \lambda_2 \in \hat{D} ,$$

$$\forall \alpha \in [0,1]$$

Luego: h es Función Cóncava. ■

04.- <u>Teorema</u> No.2 (Relación entre los Valores Optimos de la Función Dual $h(\lambda)$ y de la Función Objetivo del Programa Matemático Q)

Los valores de la Función Dual $h(\lambda)$ son una Cota InferiorDe la Función Objetivo del Programa Q : $f(X)$, $\forall\ \lambda \in D$, y $\forall\ X$ que sea solución factible del Programa Q ; es decir:

$$h(\lambda) \leq f(X) \quad , \quad \forall\ \lambda \in D \ , \ \forall\ X \in S \ / \quad g(X) \leq 0$$

independientemente de las condiciones de convexidad y diferenciabilidad del Programa Matemático Q .

<u>Demostración</u>

$$h(\lambda) = \min_{X \in S} L(X,\lambda)$$

$$= \min_{X \in S} [\ f(X) + \lambda^t g(X)\ ]$$

y por definición de mínimo de una función:

$$h(\lambda) \leq f(X) + \lambda^t g(X) \quad , \quad \forall\ X \in S \qquad \text{.....(1)}$$

Sea $\lambda \in D$ arbitrario, luego : $\lambda \geq 0$ .....(2)

Sea $X \in S$ arbitrario, tal que : $g(X) \leq 0$ .....(3)

De las condiciones (2) y (3) :

$$\lambda^t g(X) \leq 0 \quad , \ \forall\ \lambda \in D \ , \ \forall\ X \in S \ / \quad g(X) \leq 0 \qquad \text{.....(4)}$$

y de las relaciones (1) y (4) :

$$h(\lambda) \leq f(X) \quad , \quad \forall \lambda \in D \ , \ \forall X \in S \ / \ g(X) \leq 0 \qquad \blacksquare$$

NOTA:

El Teorema No.2 permite obtener una cota inferior del valor óptimo del Programa Q .

6.- Teorema No.3

Si existe una Solución Optima Factible X_o , y una Solución Dual Factible λ_o , tal que: $f(X_o) = h(\lambda_o)$ ====> X_o y λ_o son Soluciones de sus correspondientes Programas.

Demostración

Vamos a hacer la demostración por el Método de Reducción al Absurdo:

1) Supongamos que X_o no sea Solución Optima de f , osea, supongamos que X_o no sea el mínimo de f entre los vectores X factibles, entonces $\exists\ X_1$ factible, tal que:

$$f(X_1) < f(X_o) \qquad .....(i)$$

Pero por hipótesis:

$$f(X_o) = h(\lambda_o) \qquad .....(ii)$$

y de las relaciones (i) y (ii) :

$$f(X_1) < h(\lambda_o) \qquad .....(iii)$$

Pero por Teorema anterior:

$h(\lambda) \leq f(X)$, para todo par de soluciones factibles de los Programas Primal y Dual respectivamente .....(iv)

Luego de las relaciones (iii) y (iv), se tiene una contradicción.
Como la contradicción proviene de suponer que X_o no es Solución Optima del Programa Primal, debe de ocurrir lo contrario, es decir, X_o debe de ser Solución Optima del Programa Primal.

2) Análogamente, supongamos que λ_o no sea máximo de h entre todos los vectores λ factibles, entonces $\exists\ \lambda_1$ factible, tal que:

$$h(\lambda_1) > h(\lambda_o) \qquad .....(i)$$

y como por hipótesis:

$$f(X_o) = h(\lambda_o) \qquad .....(ii)$$

se tiene de las relaciones (i) y (ii) :

$$f(X_o) < h(\lambda_1) \qquad .....(iii)$$

Pero por Teorema anterior:

$h(\lambda) \leq f(X)$, para todo par de soluciones factibles de los Programas Primal y Dual respectivamente....(iv)

Luego de las relaciones (iii) y (iv) , se tiene una contradicción.

Luego: λ_o es Solución Optima del Programa Dual.

Y de las relaciones 1) y 2) el teorema esta demostrado. ∎

8. - El Programa Dual Minimáx

Definición:

Sea el Programa Matemático:

$$Q \; : \; \min Z = f(X)$$

$$\text{s. a.}$$

$$g(X) \leq 0$$

$$X \in S \subset \mathbb{R}^n$$

al que denominamos Programa Primal.

Se denomina *Programa Dual Minimáx* asociado al Programa Matemático Q, al siguiente Programa Matemático:

$$R \; : \; \max h(\lambda)$$

$$\lambda \in D \subset \mathbb{R}^m$$

el cual se expresa también de la siguiente manera:

$$R \; : \; \max_{\lambda \in D} \left[\min_{X \in S} L(X,\lambda) \right]$$

NOTA:

Los Programas: Primal Q y Dual R, se denominan un par de programas en dualidad, o un par Primal-Dual.

7.- Teorema No. 4 (Relación Fundamental entre los Programas Primal y Dual)

Teniendo en cuenta las definiciones de Programa Primal y Dual Minimáx dadas en este Capítulo, se tiene que,

El Punto (X_o, λ_o) es Punto de Silla del Lagrangeano $L(X,\lambda)$, sí y solo sí:

1) X_o es Solución Factible del Programa Primal.

2) λ_o es Solución Factible del Programa Dual.

3) $f(X_o) = h(\lambda_o)$.

Demostración

===>) Condición Necesaria

Hipótesis: (X_o, λ_o) es Punto de Silla de $L(X,\lambda)$.

Tésis: 1) $X_o \in S \quad \wedge \quad g(X_o) \leq 0$.

2) $\lambda_o \in D$.

3) $h(\lambda_o) = f(X_o)$

Como (X_o, λ_o), con $X_o \in S$ y $\lambda_o \geq 0$ es Punto de Silla de $L(X,\lambda)$, se tiene por Teorema de Kuhn-Tucker :

a) $g(X_o) \leq 0$.

b) $\lambda_o^t \, g(X_o) = 0$

c) $L(X_o, \lambda_o) = \min_{X \in S} L(X, \lambda_o)$

de la hipótesis, y de la condición a) de Kuhn-Tucker se tiene que: X_o es Solución Factible del Primal.

Análogamente, de la hipótesis, y de la condición c) de Kuhn-Tucker se tiene que: λ_o es Solución Factible del Dual, y de esta manera quedan demostradas las relaciones 1) y 2).

Resta demostrar la relación 3) :

Por definición de Función Dual se tiene:

$$
\begin{aligned}
h(\lambda_o) &= \min_{X \in S} L(X,\lambda_o) \\
&= L(X_o,\lambda_o) && \text{....(Por la relación c) de Kuhn-Tucker)} \\
&= f(X_o) + \lambda_o^t\, g(X_o) \\
&= f(X_o) && \text{....(Por la relación b) de Kuhn-Tucker)}
\end{aligned}
$$

<===) Condición Suficiente

Hipótesis: 1) X_o es Solución Factible del Programa Primal.

2) λ_o es Solución factible del Programa Dual.

3) $h(\lambda_o) = f(X_o)$

Tésis: (X_o,λ_o) es Punto de Silla de $L(X,\lambda)$.

Basados en el Teorema de Kuhn-Tucker, diremos que (X_o,λ_o) es Punto de Silla de $L(X,\lambda)$ si:

a) $g(X_o) \leq 0$

b) $\lambda_o^t\, g(X_o) = 0$

c) $L(X_o, \lambda_o) = \min_{X \in S} L(X, \lambda_o)$.

Hagamos ahora la demostración de cada una de las condiciones del Teorema de Kuhn-Tucker:

1) Por hipótesis 1) se tiene que X_o es Solución Factible del Problema Primal, es decir: $g(X_o) \leq 0$, y se cumple la condición a) del Teorema de Kuhn-Tucker.

2) Demostraremos la condición c) del teorema de Kuhn-Tucker:

$L(X_o, \lambda_o) = \min_{X \in S} L(X, \lambda_o)$.

Por el Método de reducción al Absurdo: Supongamos que $L(X_o, \lambda_o)$ no minimiza $L(X, \lambda_o)$; es decir existe otro vector $X_1 \in S$ tal que:

$$h(\lambda_o) = \min_{X \in S} L(X, \lambda_o) = L(X_1, \lambda_o) < L(X_o, \lambda_o)$$

ó :

$$f(X_1) + \lambda_o^t g(X_1) < f(X_o) + \lambda_o^t g(X_o) \qquad \text{.....(2.1)}$$

Por Hipótesis 3) :

$$f(X_o) = h(\lambda_o)$$

$$= f(X_1) + \lambda_o^t g(X_1) \quad \text{...(Por Hipótesis Auxiliar)..(2.2)}$$

Reemplazando la condición (2.2) en la condición (2.1) , y simplificando, se obtiene:

$$\lambda_o^t g(X_o) > 0 \qquad \text{.....(2.3)}$$

Pero por hipótesis 1) y 2) : $\lambda_o \geq 0$ y $g(X_o) \leq 0$
luego:

$$\lambda_o^t\, g(X_o) \leq 0 \qquad \ldots\ldots(2.4)$$

y de las relaciones (2.3) y (2.4) se tiene una contradicción, como la contradicción proviene de suponer que $L(X_o,\lambda_o)$ no minimiza $L(X,\lambda_o)$, debe de ocurrir lo contrario, es decir, $L(X_o,\lambda_o)$ minimiza $L(X,\lambda_o)$.

3) Resta demostrar la condición b) del teorema de Kuhn-Tucker.

Por definición de Función Dual:

$$h(\lambda) = \min_{X \in S} L(X,\lambda)$$

y :

$$h(\lambda_o) = \min_{X \in S} L(X,\lambda_o)$$

$$= L(X_o,\lambda_o) \quad \text{(Por condición c) de Kuhn-Tucker)}$$

$$= f(X_o) + \lambda_o^t\, g(X_o) \qquad \ldots\ldots(3.1)$$

Por hipótesis 3) :

$$h(\lambda_o) = f(X_o) \qquad \ldots\ldots(3.2)$$

de las condiciones (3.1) y (3.2) :

$$\lambda_o^t\, g(X_o) = 0$$

y la condición c) del teorema de Kuhn-Tucker esta demostrada.

De los pasos 1) , 2) y 3) , y del Teorema de Kuhn-Tucker:

(X_o, λ_o) es Punto de Silla de $L(X, \lambda)$. ■

NOTAS:

1) El Teorema que se acaba de demostrar, garantiza que (X_o, λ_o) es Punto de Silla del Lagrangeano $L(X, \lambda)$, si X_o y λ_o son Soluciones Factibles de sus respectivos Programas, y $f(X_o) = h(\lambda_o)$, en cuyo caso X_o es Solución Optima de Programa Primal.

2) La busqueda de la Componente Principal X_o , considerada como vector que minimiza $L(X, \lambda_o)$ con respecto a $X \in S$, tropieza con la dificultad de que λ_o es desconocido a priori, de ahí que el mínimo de $L(X, \lambda)$ con respecto a $X \in S$ se estudia en función de λ , considerando todos los $\lambda \geq 0$.

3) En los Programas cuyo Lagrangeano posee un Punto de Silla, coinciden los valores óptimo de los Programas Primal y Dual.

4) Si no existe Punto de Silla, no coinciden los valores óptimos de ambos programas, y el Teorema no permite localizar la Solución Óptima del Problema Primal, lo cual no quiere decir que ésta no exista, ya que las condiciones de Punto de Silla son condiciones suficientes, pero no necesarias de optimalidad.

3.- Aplicaciones

I.- Resuelva el siguiente Programa Matemático:

$$T \; : \quad \min Z = (X_1 - 1)^2 + (X_2 - 2)^2$$

s. a.

$$X_2 - X_1 \leq 1$$

$$X_1 + X_2 \leq 2$$

$$X_1 \, , \, X_2 \geq 0$$

usando el concepto de Dual Minimáx.

Solución

1) El Lagrangeano asociado al Programa Matemático es :

$$L(X,\lambda) = (X_1 - 1)^2 + (X_2 - 2)^2 + \lambda_1 (X_2 - X_1 - 1) + \lambda_2 (X_1 + X_2 - 2)$$

donde $\lambda_1 \, , \, \lambda_2 \geq 0$.

Reordenando términos:

$$L(X,\lambda) = (X_1 - 1)^2 + (X_2 - 2)^2 + (\lambda_2 - \lambda_1) X_1 + (\lambda_1 + \lambda_2) X_2 - (\lambda_1 + 2 \lambda_2)$$

2) Hallando el mínimo de $L(X,\lambda)$ con respecto a $X = (X_1, X_2)$, se tiene:

$$2.1) \quad \frac{\partial L(X,\lambda)}{\partial X_1} = 2 (X_1 - 1) + \lambda_2 - \lambda_1 = 0 \implies X_1^0 = 1 + \frac{1}{2} (\lambda_1 - \lambda_2)$$

$$\frac{\partial L(X,\lambda)}{\partial X_2} = 2(X_2 - 2) + \lambda_1 + \lambda_2 = 0 \implies X_2^0 = 2 - \frac{1}{2}(\lambda_1 + \lambda_2)$$

2.2) $$\frac{\partial^2 L(X,\lambda)}{\partial X_1^2} = 2$$

$$\frac{\partial^2 L(X,\lambda)}{\partial X_2^2} = 2$$

$$\frac{\partial^2 L(X,\lambda)}{\partial X_1 \partial X_2} = \frac{\partial}{\partial X_1}\left[\frac{\partial L(X,\lambda)}{\partial X_2}\right]$$

$$= \frac{\partial}{\partial X_1}\left[2(X_2 - 2) + \lambda_1 + \lambda_2\right]$$

$$= 0$$

2.3) Como:

$$\Delta = \left[\frac{\partial^2 L(X_1^0, X_2^0, \lambda)}{\partial X_1^2}\right]\left[\frac{\partial^2 L(X_1^0, X_2^0, \lambda)}{\partial X_2^2}\right] - \left[\frac{\partial^2 L(X_1^0, X_2^0, \lambda)}{\partial X_1 \partial X_2}\right]^2$$

$$= 4 > 0$$

y :

a) $$\frac{\partial^2 L(X_1^o, X_2^o, \lambda)}{\partial X_1^2} = 2 > 0$$

b) $$\frac{\partial^2 L(X_1^o, X_2^o, \lambda)}{\partial X_2^2} = 2 > 0$$

Luego:

$L(X_1^o, X_2^o, \lambda)$ toma su valor mínimo en el Punto

$$(X_1^o, X_2^o) = \left(1 + \tfrac{1}{2}(\lambda_1 - \lambda_2) \;,\; 2 - \tfrac{1}{2}(\lambda_1 + \lambda_2)\right)$$

3) Reemplazando los valores de X_1^o y X_2^o obtenidos en el paso anterior en el Lagrangeano, se obtiene la Función Dual $h(\lambda)$:

$$h(\lambda) = \min_{X \in S} L(X, \lambda)$$

$$= (X_1^o - 1)^2 + (X_2^o - 2)^2 + \lambda_1 (X_2^o - X_1^o - 1) + \lambda_2 (X_1^o + X_2^o - 2)$$

$$= \left[1 + \tfrac{1}{2}(\lambda_1 - \lambda_2) - 1\right]^2 + \left[2 - \tfrac{1}{2}(\lambda_1 + \lambda_2) - 2\right]^2$$

$$+ \lambda_1 \left[2 - \tfrac{1}{2}(\lambda_1 + \lambda_2) - 1 - \tfrac{1}{2}(\lambda_1 - \lambda_2) - 1\right]$$

$$+ \lambda_2 \left[1 + \tfrac{1}{2}(\lambda_1 - \lambda_2) + 2 - \tfrac{1}{2}(\lambda_1 + \lambda_2) - 2\right]$$

$$= \frac{1}{4}(\lambda_1 - \lambda_2)^2 + \frac{1}{4}(\lambda_1 + \lambda_2)^2 + \lambda_1(-\lambda_1) + \lambda_2(1 - \lambda_2)$$

$$= \frac{1}{4}(\lambda_1^2 - 2\lambda_1\lambda_2 + \lambda_2^2) + \frac{1}{4}(\lambda_1^2 + 2\lambda_1\lambda_2 + \lambda_2^2) - \lambda_1^2 + \lambda_2 - \lambda_2^2$$

$$= -\frac{1}{2}\lambda_1^2 - \frac{1}{2}\lambda_2^2 + \lambda_2$$

4) Hallando el dominio de definición de la Función Dual:

Para que la Solución hallada en el Paso 2) sea factible, se debe de cumplir que:

$$X_o = (X_1^o, X_2^o) \geq 0$$

$$X_1^o = 1 + \frac{1}{2}(\lambda_1 - \lambda_2) \geq 0 \quad ===> \quad \lambda_1 - \lambda_2 \geq -2$$

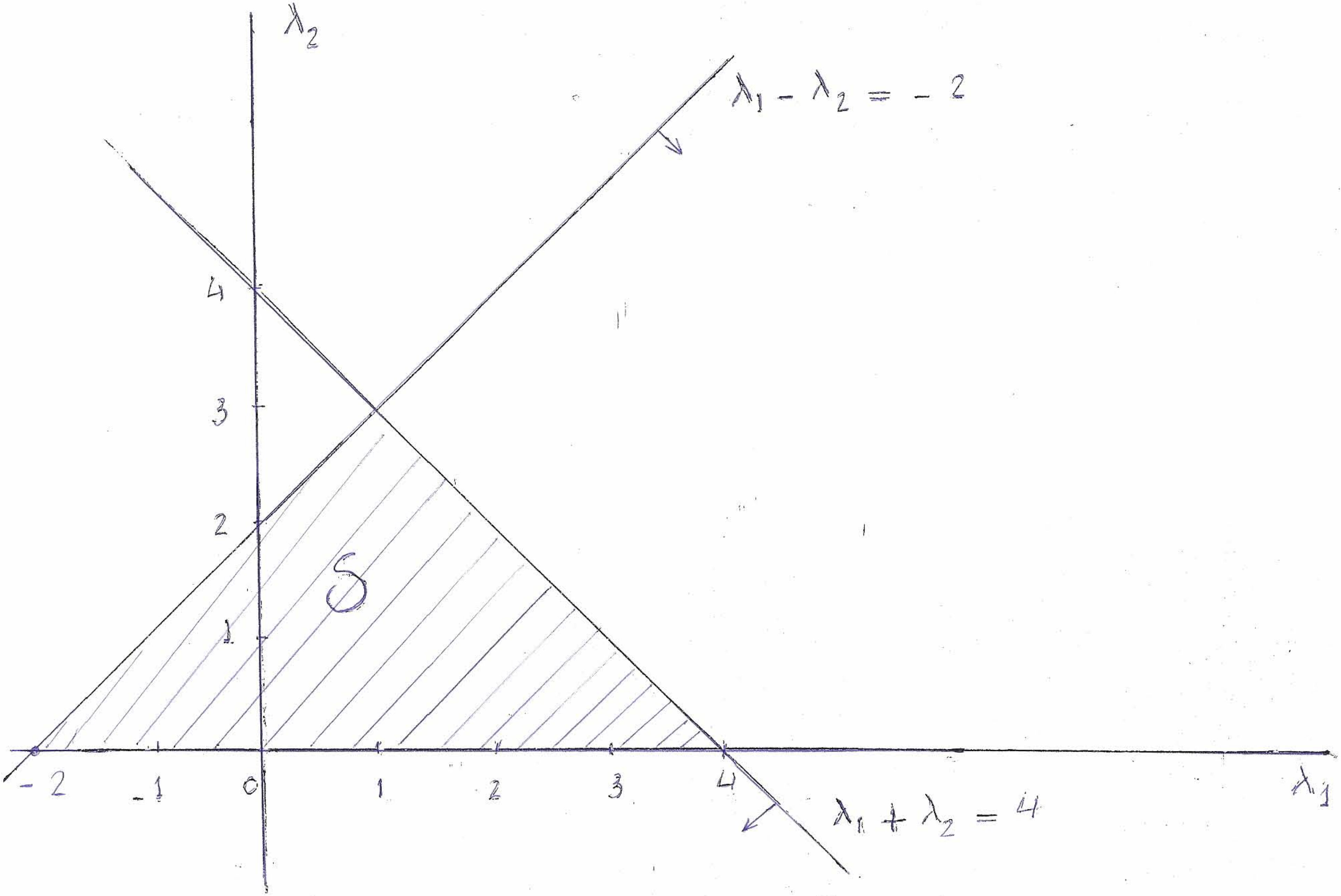
$$X_2^o = 2 - \frac{1}{2}(\lambda_1 + \lambda_2) \geq 0 \quad ===> \quad \lambda_1 + \lambda_2 \leq 4$$

y el dominio de la Función Dual $h(\lambda_1, \lambda_2)$, de acuerdo a los anteriores resultados es:

$$D = \left\{(\lambda_1, \lambda_2) \in \mathbb{R}^2 \;/\; \lambda_1 - \lambda_2 \geq -2 \;,\; \lambda_1 + \lambda_2 \leq 4 \;,\; \lambda_2 \geq 0\right\}$$

cuya gráfica se representa en el Gráfido No.3 .

GRAFICO No. 5


λ_2
$\lambda_1 - \lambda_2 = -2$
4
3
2
1
S
-2
-1
0
1
2
3
4
$\lambda_1 + \lambda_2 = 4$
λ_1

5) Hallemos ahora: $\max_{\lambda \in D \subset \mathbb{R}^2} h(\lambda) = \max_{\lambda \in D} \left[\min_{X \in S} L(X,\lambda) \right]$.

Tenemos:

$$h(\lambda_1,\lambda_2) = -\frac{1}{2}\lambda_1^2 - \frac{1}{2}\lambda_2^2 + \lambda_2$$

luego:

$$\frac{\partial h(\lambda_1,\lambda_2)}{\partial \lambda_1} = -\lambda_1 = 0 \implies \lambda_1 = 0$$

$$\frac{\partial h(\lambda_1,\lambda_2)}{\partial \lambda_2} = -\lambda_2 + 1 = 0 \implies \lambda_2 = 1$$

de donde:

$$(\lambda_1^0,\lambda_2^0) = (0,1)$$

y :

$$h(\lambda_1^0,\lambda_2^0) = \frac{1}{2} .$$

6) Como el conjunto:

$$D = \left\{(\lambda_1,\lambda_2) \in \mathbb{R}^2 \ / \ \lambda_1 - \lambda_2 \geq -2 \ , \ \lambda_1 + \lambda_2 \leq 4 \ , \ \lambda_2 \geq 0 \right\}$$

es convexo, se tiene por Teorema No.1 que:

$$h(\lambda_1,\lambda_2) = -\frac{1}{2}\lambda_1^2 - \frac{1}{2}\lambda_2^2 + \lambda_2$$

es una función cóncava.

7) Resta justificar que $h(\lambda_1,\lambda_2)$ tiene máximo absoluto en $(\lambda_1^0,\lambda_2^0) = (0,1)$.

Tenemos:

7.1) $$\frac{\partial\, h(\lambda_1,\lambda_2)}{\partial\, \lambda_1} = -\lambda_1$$

$$\frac{\partial\, h(\lambda_1,\lambda_2)}{\partial\, \lambda_2} = -\lambda_2 + 1$$

7.2) $$\frac{\partial^2\, h(\lambda_1,\lambda_2)}{\partial\, \lambda_1^2} = -1$$

$$\frac{\partial^2\, h(\lambda_1,\lambda_2)}{\partial\, \lambda_2^2} = -1$$

7.3) $$\frac{\partial^2\, h(\lambda_1,\lambda_2)}{\partial\, \lambda_1 \quad \partial\, \lambda_2} = 0$$

7.4) Como:

$$\Delta = \left[\frac{\partial^2 h(\lambda_1^0, \lambda_2^0)}{\partial \lambda_1^2} \right] \left[\frac{\partial^2 h(\lambda_1^0, \lambda_2^0)}{\partial \lambda_2^2} \right] - \left[\frac{\partial^2 h(\lambda_1^0, \lambda_2^0)}{\partial \lambda_1 \quad \partial \lambda_2} \right]^2$$

$$= (-1) \quad (-1) \quad - \quad 0$$

$$= 1 \quad \geq \quad 0$$

y :

a) $$\frac{\partial^2 h(\lambda_1^0, \lambda_2^0)}{\partial \lambda_1^2} = -1 \quad < \quad 0$$

b) $$\frac{\partial^2 h(\lambda_1^0, \lambda_2^0)}{\partial \lambda_2^2} = -1 \quad < \quad 0$$

luego: $h(\lambda_1, \bar{\lambda}_2)$ tiene un máximo en $(\lambda_1^0, \lambda_2^0) = (0,1)$, como además $h(\lambda_1, \lambda_2)$ es una función cóncava, luego: $h(\lambda_1, \lambda_2)$ tiene un máximo absoluto en $(\lambda_1^0, \lambda_2^0) = (0,1)$.

8) Vamos ahora a hallar la Solución Optima del Programa Matemático T .

Tenemos:

8.1) $X_0 = (X_1^0, X_2^0) = \left(\frac{1}{2} , \frac{3}{2} \right)$, solución factible del Problema T.

8.2) $\lambda_o = (\lambda_1^o, \lambda_2^o) = (0,1)$, solución factible del Problema Dual.

8.3) $f(X_o) = f(\frac{1}{2}, \frac{3}{2}) = (\frac{1}{2} - 1)^2 + (\frac{3}{2} - 2)^2 = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$.

$h(\lambda_o) = h(0,1) = -\frac{1}{2} + 1 = \frac{1}{2}$

es decir: $f(X_o) = h(\lambda_o)$.

Y por el Teorema No.4 :

$(X_o, \lambda_o) = (\frac{1}{2}, \frac{3}{2}, 0, 1)$ es Punto de Silla de $L(X_o, \lambda_o)$.

Y por el Teorema de Kuhn-Tucker:

$X_o = (\frac{1}{2}, \frac{3}{2})$ es Solución Optima del Problema T . ■

II.- Resuelva el siguiente Programa Matemático:

$$P \ : \ \min Z = 4 X_1^2 + 2 X_1 X_2 + X_2^2$$

s. a.

$$3 X_1 + X_2 \geq 6$$

$$X_1 \geq 0$$

$$X_2 \geq 0$$

usando el concepto de Dual Minimáx.

<u>Solución</u>

1) El Lagrangeano asociado al Programa Matemático es :

$$L(X,\lambda) = 4 X_1^2 + 2 X_1 X_2 + X_2^2 + \lambda (6 - 3 X_1 - X_2)$$

donde $\lambda \geq 0$.

2) Hallando el mínimo de $L(X,\lambda)$ con respecto a $X = (X_1, X_2)$ se tiene:

2.1) $$\frac{\partial L(X,\lambda)}{\partial X_1} = 8 X_1 + X_2 - 3 \lambda = 0$$

$$\frac{\partial L(X,\lambda)}{\partial X_2} = 2 X_1 + 2 X_2 - \lambda = 0$$

resolviendo este sistema en función del parámetro λ se tiene:

$$X_1^0 = \frac{1}{3} \lambda$$

$$\frac{\partial^2 L(X_1^0, X_2^0, \lambda)}{\partial X_1^2} = 8 > 0$$

$$\frac{\partial^2 L(X_1^0, X_2^0, \lambda)}{\partial X_2^2} = 2 > 0$$

Luego, $L(X_1, X_2, \lambda)$ toma su valor mínimo en el Punto :

$$(X_1^0, X_2^0) = \left(\frac{1}{3} \lambda \, , \, \frac{1}{6} \lambda \right)$$

Reemplazando los valores de X_1^0 y X_2^0 en el Lagrangeano, se obtiene la Función Dual :

$$h(\lambda) = \min_{X \in S} L(X, \lambda)$$

$$= 4 \left(\frac{1}{3} \lambda \right)^2 + 2\left(\frac{1}{3} \lambda \right)\left(\frac{1}{6} \lambda \right) + \left(\frac{1}{6} \lambda \right)^2 + \lambda \left(6 - \lambda - \frac{1}{6} \lambda \right)$$

$$= \frac{4}{9} \lambda^2 + \frac{1}{9} \lambda^2 + \frac{1}{36} \lambda^2 + 6 \lambda - 6 \lambda^2 - \frac{1}{6} \lambda^2$$

$$= \left(\frac{4}{9} + \frac{1}{9} + \frac{1}{36} - \frac{1}{6} - 6 \right) \lambda^2 + 6 \lambda$$

$$= 6 \lambda - \frac{21}{36} \lambda^2$$

$$= 6 \lambda - \frac{7}{12} \lambda^2$$

donde el dominio de definición de la función dual h es la semirecta real positiva, pues $\forall\ \lambda \geq 0$ existe Solución al Problema: $\min\limits_{X \in S} L(X,\lambda)$.

3) Hallemos ahora: $\max\limits_{\lambda \geq 0} h(\lambda) = \max\limits_{\lambda \geq 0} \left[\min\limits_{X \in S} L(X,\lambda) \right]$

como:

$$h(\lambda) = 6\lambda - \frac{7}{12}\lambda^2$$

luego:

$$\frac{\partial\, h(\lambda)}{\partial \lambda} = 6 - \frac{14}{12}\lambda = 0 \quad ===> \quad \lambda = \frac{72}{14} = \frac{36}{7}$$

$$\frac{\partial^2\, h(\lambda)}{\partial \lambda^2} = -\frac{14}{12} < 0$$

===> la función dual h tiene un máximo en $\lambda_o = \frac{36}{7}$, y su valor máximo es:

$$h(\lambda_o) = h\left(\frac{36}{7}\right) = 6\left(\frac{36}{7}\right) - \frac{7}{12}\left(\frac{36}{7}\right)^2 = \frac{36}{7}\left(6 - \frac{7}{12}\cdot\frac{36}{7}\right)$$

$$= \frac{36}{7}(6-3) = 3\cdot\frac{36}{7} \approx 15.428$$

4) A modo de ilustración hallaremos la Solución Optima del Programa Matemático P .

Como: $\lambda_o = \frac{36}{7}$

luego: $X_1^o = \frac{1}{3}\left(\frac{36}{7}\right) = \frac{12}{7}$

$X_2^o = \frac{1}{6}\left(\frac{36}{7}\right) = \frac{6}{7}$

y : $X^o = (X_1^o, X_2^o) = \left(\frac{12}{7}, \frac{6}{7}\right)$ es la Solución factible del Programa.

Además:

$$f(X^o) = f\left(\frac{12}{7}, \frac{6}{7}\right) = 4\left(\frac{12}{7}\right)^2 + 2\left(\frac{12}{7}\right)\left(\frac{6}{7}\right) + \left(\frac{6}{7}\right)^2$$

$$= 2.\frac{12}{7}\left(2.\frac{12}{7} + \frac{6}{7}\right) + \left(\frac{6}{7}\right)^2$$

$$= \frac{24}{7}\left(\frac{30}{7}\right) + \left(\frac{6}{7}\right)^2$$

$$= \frac{720}{49} + \frac{36}{49}$$

$$= \frac{756}{49} = \frac{108}{7} \approx 15.428$$

5) Tenemos que:

$X_o = (X_1^o, X_2^o) = \left(\frac{12}{7}, \frac{6}{7}\right)$ es Solución factible del Programa P

$\lambda_o = \frac{36}{7}$ es la Solución factible del Programa Dual.

$f(X_o) = h(\lambda_o)$

y por Teorema No. 4 :

$(X_o, \lambda_o) = \left(\frac{12}{7}, \frac{6}{7}, \frac{36}{7}\right)$ es Punto de Silla de $L(X, \lambda)$.

Y por Teorema de Kuhn-Tucker:

$X_o = \left(\frac{12}{7}, \frac{6}{7}\right)$ es Solución Optima del Problema P .

CONCLUSIONES

1.- Las Condiciones de Kuhn y Tucker aplicadas a un Programa Matemático de la forma:

$$P \ : \ \min Z = f(X)$$
$$\text{s.a.}$$
$$g_i(X) \leq 0 \quad , \quad i = 1,2,\ldots,m.$$
$$X \geq 0$$

son condiciones necesarias y suficientes de optimalidad global, en caso de que el Programa Matemático sea convexo y diferenciable, ya que tales condiciones permiten indentificar si el Punto (X_o, Y_o) es un punto de silla del Lagrangeano del Programa, de lo que se deduce la optimalidad global de X_o, de acuerdo con los Teoremas anteriores.

2.- En condiciones de diferenciabilidad y convexidad del Programa P , coinciden los conjuntos de:

a) Soluciones Optimas del Programa.

b) Puntos de Silla del Lagrangeano.

c) Soluciones de las Condiciones de Kuhn y Tucker.

3.- Si el Programa Matemático P es diferenciable pero no convexo, las Condiciones de Kuhn y Tucker constituyen condiciones necesarias pero no suficientes de optimalidad, pues existen puntos que la verifican y no son Soluciones Optimas del Programa (Ver el ejemplo en el Capítulo respectivo).

4.- Mediante el concepto de Función Dual h :

$$h(\lambda) = \min_{X \in S} L(X,\lambda)$$

se estudia el mínimo de Lagrangeano con respecto a $X \in S$, para todo los $\lambda \geq 0$.

5.- El Teorema de relación fundamental entre los Programas Primal y Dual establece las condiciones bajo las cuales el Lagrangeano $L(X,\lambda)$ posee Punto de Silla.

B I B L I O G R A F I A

1.- Duality in Non Linear Programing.
Balinski and Baumol.
Review of Economic Estudies, Junio 1968.

2.- Non Linear Programing.
Bazaraa - Shetty.
John Wiley, New York, 1977.

3.- Non Linear Programing.
Mangasarian O. L.
Mc. Graw Hill, New York 1969.

4.- Mathematical Programing.
Vajda S.
Addison Wesley, 1961.

5.- Optimization Theory for Larges Sistems.
Lasdon L.
Mc Millan, 1970.